(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 831 458 A2

# (12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
**25.03.1998 Bulletin 1998/13**

(51) Int. Cl.[6]: **G10L 5/02**, G10L 7/02, G10L 9/06, G10L 9/18

(21) Application number: **97116245.8**

(22) Date of filing: **18.09.1997**

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE**
Designated Extension States:
**AL LT LV RO SI**

(30) Priority: **18.09.1996 JP 246726/96**
**13.03.1997 JP 76668/97**
**13.03.1997 JP 76672/97**
**13.03.1997 JP 76682/97**
**13.03.1997 JP 76693/97**
**13.03.1997 JP 76695/97**

(71) Applicant:
**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**
**Shinjuku-ku, Tokyo 163-19 (JP)**

(72) Inventors:
- **Aoki, Mariko**
  **Yokohama-shi, Kanagawa 227 (JP)**
- **Aoki, Shigeaki**
  **Yokosuka-shi, Kanagawa 238 (JP)**
- **Matsui, Hiroyuki**
  **Yokohama-shi, Kanagawa 235 (JP)**
- **Nishino, Yutaka**
  **Miura-shi, Kanagawa 238-01 (JP)**
- **Okamoto, Manabu**
  **Yokohama-shi, Kanagawa 236 (JP)**

(74) Representative:
**Hoffmann, Eckart, Dipl.-Ing.**
**Patentanwalt,**
**Bahnhofstrasse 103**
**82166 Gräfelfing (DE)**

## (54) Method and apparatus for separation of sound source, program recorded medium therefor, method and apparatus for detection of sound source zone; and program recorded medium therefor

(57) A time difference $\Delta\tau$ between the arrival of acoustic signals from sound sources to microphones 1, 2 is detected from output channel signals L, R from microphones 1, 2. By Fourier transform, the signals L, R are divided into respective frequency bands L(f1) - L(fn), R(f1) - R(fn). Differences $\Delta\tau_i$ ( i = 1, 2, • • • n ) in the time-of-arrival of L(f1) - L(fn) and R(f1) - R(fn) to the microphones 1, 2 as well as a signal level difference $\Delta Li$ are detected. L(f1) - L(fn), R(f1) - R(fn) are divided into a low range of $fi < 1/(2\Delta\tau)$, a middle range of $1/(2\Delta\tau) < fi < 1/\Delta\tau$, and a high range of $fi > 1/\Delta\tau$. Utilizing $\Delta\tau_i$ for the low range, $\Delta Li$ and $\Delta\tau_i$ for the middle range and $\Delta Li$ for the high range, a determination is made from which sound source L(fi), R(fi) are oncoming to deliver outputs separately for each sound source. The outputs are subject to an inverse Fourier transform for synthesis separately for each sound source.

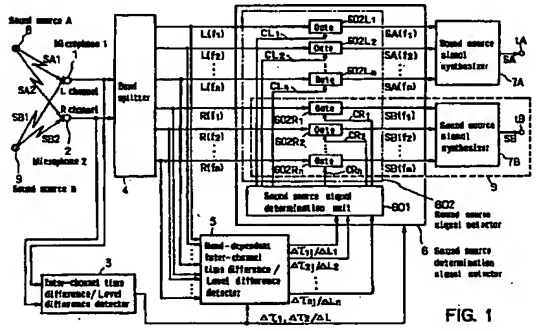


FIG. 1

EP 0 831 458 A2

## Description

Background of the Invention :

The invention relates to a method of separating/extracting a signal of at least one sound source from a complex signal comprising a mixture of a plurality of acoustic signals produced by a plurality of sound sources such as voice signal sources and various environmental noise sources, an apparatus for separating sound source which is used in implementing the method, and recorded medium having a program recorded therein which is used to carry out the method in a computer.

An apparatus for separating sound source of the kind described is used in a variety of applications including a sound collector used in a television conference system, a sound collector used for transmission of a voice signal uttered in a noisy environment, or a sound collector in a system which distinguishes between the types of sound sources, for example :

A conventional technology for separating sound source comprises estimating fundamental frequencies of various signals in the frequency domain, extracting harmonics structures, and collecting components from a signal source for synthesis.

However, the technology suffers from (1) the problem that signals which permit such a separation are limited to those having harmonic structures which resemble the harmonic structures of vowel sounds of voices or musical tones; (2) the difficulty of separating sound sources from each other in real time because the estimation of the fundamental frequencies generally require an increased length of time for processing; and (3) the insufficient accuracy of separation which results from erroneous estimations of harmonic structures which cause frequency components from other sound sources to be mixed with the extracted signal and cause such components to be perceived as noise.

A conventional sound collector in a communication system also suffers from the howling effect that a voice reproduced by a loudspeaker on the remote end is mixed with a voice on the collector side. A howling suppression in the art includes a technique of suppressing of the unnecessary components from the estimation of the harmonic structures of the signal to be collected and a technique of defining a microphone array having a directivity which is directed to a sound source from which a collection is to be made.

The former technique is effective only when the signal has a high pitch response while signals to be suppressed have a flat frequency response as a consequence of utilizing the harmonic structures. Thus, the howling suppression effect is reduced in a communication system in which both the sound source from which a collection is desired and the remote end source deliver a voice. The latter technique of using the microphone array requires an increased number of microphones to achieve a satisfactory detectivity, and accordingly, it is difficult to use a compact arrangement. In addition, if the directivity is enhanced, a movement of the sound source results in an extreme degradation in the performance, with concominant reduction in howling suppression effect.

As a technique of detecting a zone in which a sound source uttering a voice or speaking source is located in a space in which a plurality of sound sources are disposed, a technique is known in the art which uses a plurality of microphones and detects the location of the sound source from differences in the time required for an acoustic signal from the source to reach individual microphones. This technique utilizes a peak value of cross-correlation between output voice signals from the microphones to determine a difference in time required for the acoustic signal to reach each microphone, thus detecting the location of the sound source.

Unfortunately, this detection technique requires an increased length of time for calculation of cross-correlation functions which must be performed by additions and multiplications of a data length which is twice the data length read already.

The use of a histogram is effective in detecting a peak among the cross-correlations. However, a histogram formed on a time axis causes a time delay. To provide a histogram without causing a time delay, it is contemplated to divide the signal into bands, and to form a histogram over all the bands. However, it is necessary to employ a signal having a bandwidth greater than a given value to form a cross-correlation function, and accordingly, the division of the signal is limited to several bands at most. Hence, the histogram must be formed on the time axis using a signal having a certain length, but it is difficult with this technique to detect the location of the sound source in real time.

An estimation of direction of a sound source by a processing technique in which outputs from a pair of microphones are each divided into a plurality of bands is disclosed in Japanese Laid-Open Patent Application Number 87, 903 / 93. The disclosed technique requires a calculation of a cross-correlation between signals in corresponding divided bands, and hence suffers from an increased length of processing time.

It is an object of the invention to provide a method and an apparatus which separates / extracts an acoustic signal from a sound source that does not have a harmonic structure, and thus enables a separation of a sound source without dependence on the variety of the sound source and enables such a separation in real time, and a program recorded medium therefor.

It is another object of the invention to provide a method and an apparatus for the separation of a sound source with



BNSDOCID: <EP_____0831458A2_I_>

a high accuracy and with a reduced level of noise, and a program recorded medium therefor.

It is a further object of the invention to provide a method and an apparatus for separation of a sound source which permits the howling to be suppressed to a sufficiently low level for any signal, and a program recorded medium therefor.

It is still another object of the invention to provide a method and an apparatus for detection of a sound source zone in real time, and a program recorded medium therefor.

## SUMMARY OF THE INVENTION :

In accordance with the invention, a method of separating a sound source comprises the steps of

providing a plurality of microphones which are located as separated from each other, each microphone providing an output channel signal which is divided into a plurality of frequency bands in a frequency division process such that essentially and principally a signal component from a single sound source resides in each band;
detecting, for each common band of respective output channel signals, a difference in a parameter such as a level (power) and / or time of arrival (phase) of an acoustic signal reaching each microphone which undergoes a change attributable to the locations of the plurality of microphones as a band-dependent inter-channel parameter value difference;
on the basis of the band-dependent inter-channel parameter value differences for each band, determining in a sound source signal determination process which one of the respective band-divided output channel signals for a particular band comes from which one of the sound sources;
on the basis of a determination rendered in the sound source signal determination process, selecting in a sound source signal selection process at least one of the signals coming from a common sound source from the band-divided output signals;
and synthesizing in a sound source synthesis process a plurality of band signals selected as signals from a common sound source in the sound source signals selection process into a sound source signal.

In an embodiment of the invention, the band-dependent levels of the respective output channel signals which are divided in the band division process are detected. The band-dependent levels for a common band are compared between channels, and based on the results of such a comparison, a sound source ( or sources ) which is not uttering a voice is detected. A detection signal corresponding to the sound source which is not uttering a voice is used to suppress a synthesized signal corresponding to the sound source which is not uttering a voice from among the sound sources signal which are synthesized in the sound source synthesis process.

In another embodiment of the invention, differences in the time required for the respective output channel signals which are divided in the band division process to reach respective microphones are detected for each common band. The band-dependent differences in time thus detected for each common band are compared between the channels, and on the basis of the results of such a comparison, a sound source (or sources) which is not uttering a voice is detected. A detection signal corresponding to the sound source which is not uttering a voice is used to suppress a synthesized signal corresponding to the sound source which is not uttering a voice from among the sound source signals which are synthesized in the sound source synthesis process.

In a further embodiment of the invention, at least one of the sound sources is a speaker, and at least one of the other sound sources is electroacoustical transducer means which transduces a received signal oncoming from the remote end into an acoustic signal. The sound source signal selection process interrupts components in the band-divided channel signals which belong to the acoustic signal from the electracoustical transducer means, and selects components of the voice signal form the speaker. The sound source signal synthesized in the sound source synthesis process is transmitted to the remote end.

In accordance with the invention, a method of detecting a sound source zone comprises providing a plurality of microphones which are located as separated from each other, each microphone providing an output channel signal which is divided into a plurality of frequency bands such that essentially and principally a signal component from a single sound source resides in each band, detecting, for each common band of respective output channel signals, a difference in a parameter such as a level (power) and / or time of arrival (phase) of the acoustic signal reaching each microphone which undergoes a change attributable to the locations of the plurality of microphone, comparing the parameter values thus detected for each band between the channels, and on the basis of the result of such comparison, determining a zone in which the sound source of the acoustic signal reaching the microphone is located.

## BRIEF DESCRIPTION OF THE DRAWINGS :

Fig. 1 is a functional block diagram of an apparatus for separation of sound source according to an embodiment of the invention;



BNSDOCID: <EP_____0831458A2_I_>

Fig. 2 is a flow diagram illustrating a processing procedure used in a method of separating a sound source according to an embodiment of the invention;

Fig. 3 is a flow diagram of an exemplary processing procedure for determining inter-channel time differences $\Delta\tau_1$, $\Delta\tau_2$ shown in Fig. 2;

Figs. 4 A and B are diagrams showing examples of the spectrums for two sound source signals;

Fig. 5 is a flow diagram illustrating a processing procedure in a method of separating sound source according to an embodiment of the invention in which the separation takes place by utilizing inter-channel level differences;

Fig. 6 is a flow diagram showing a part of a processing procedure according to the method of separating a sound source according to the embodiment of the invention in which both inter-channel level differences and inter-channel time-of-arrival differences are utilized;

Fig. 7 is a flow diagram which continues to step S08 shown in Fig. 6;

Fig. 8 is a flow diagram which continues to step S09 shown in Fig. 6;

Fig. 9 is a flow diagram which continues to step S10 shown in Fig. 6 and which also continues to steps S20 and S30 shown in Fig. 7 and 8, respectively;

Fig. 10 is a functional block diagram of an embodiment in which sound source signals of different frequency bands are separated from each other;

Fig. 11 is a functional block diagram of an apparatus for separation of sound source according to another embodiment of the invention in which an arrangement is added to suppress unnecessary sound source signal utilizing a level difference;

Fig. 12 is a schematic illustration of the layout of three microphones, their coverage zones and two sound sources;

Fig. 13 is a flow diagram illustrating an exemplary procedure of detecting a sound source zone and generating a suppression control signal when only one sound source is uttering a voice;

Fig. 14 is a schematic illustration of the layout of three microphones, their coverage zones and three sound sources;

Fig. 15 is a flow diagram illustrating a procedure of detecting a zone for a sound source which is uttering a voice and generating a suppression control signal where there are three sound sources;

Fig. 16 is a schematic illustration of the layout in which three microphones are used to divide the space into three zones, also illustrating the layout of sound sources;

Fig.17 is a flow diagram illustrating a processing procedure used in an apparatus for separating the sound source according to the invention for generating a control signal which is used to suppress a synthesized sound source signal for a sound source which is not uttering a voice;

Fig. 18 is a functional block diagram of an apparatus for separating a sound source according to another embodiment of the invention in which an arrangement is added for suppressing unnecessary sound source signal by utilizing a time-of-arrival difference;

Fig. 19 is a schematic illustration of an exemplary relationship between a speaker, a loudspeaker and a microphone in an apparatus for separating a sound source according to the invention which is applied to the suppression of runaround sound;

Fig.20 is a functional block diagram of an apparatus for separating a sound source according to a further embodiment of the invention which is applied to the suppression of runaround sound;

Fig. 21 is a functional block diagram of part of an apparatus for separating a sound source according to still another embodiment of the invention which is applied to the suppression of runaround sound;

Fig. 22 is a functional block diagram of an apparatus for separating a sound source according to an embodiment of the invention in which a division into bands takes place after a power spectrum is determined;

Fig. 23 is a functional block diagram of an apparatus for zone detection according to an embodiment of the invention;

Fig. 24 is a flow diagram illustrating a processing procedure used in the zone detecting method according to the embodiment of the invention;

Fig. 25 is a chart showing the varieties of sound sources used in an experiment for the invention;

Fig. 26 is a diagram illustrating voice spectrums before and after processing according to the method of embodiments shown in Figs. 6 to 9;

Fig. 27 are diagrams showing results of a subjective evaluation experiment which uses the method of embodiment shown in Figs. 6 to 9;

Fig. 28 shows voice waveforms after the processing according to the method of embodiments shown in Figs. 6 to 9 together with the original voice waveform;

Fig. 29 shows results of experiments conducted for the method of separating a sound source as illustrated in Figs. 6 to 9 and the apparatus for separating sound source shown in Fig. 11; and

Fig. 30 is a functional block diagram of another embodiment of the invention which is applied to the suppression of runaround sound.



BNSDOCID: <EP_____0831458A2_I_>

DESCRIPTION OF PREFERRED EMBODIMENTS

Fig. 1 shows an embodiment of the invention. A pair of microphones 1 and 2 are disposed at a spacing from each other, which may be on the order of 20 cm, for example, for collecting acoustic signals from the sound sources A, B and converting them into electrical signals. An output from the microphone 1 is referred to as an L channel signal, and an output form the microphone 2 is referred to as an R channel signal. Both the L channel and the R channel signal are fed to an inter-channel time difference / level difference detector 3 and a bandsplitter 4. In the bandsplitter 4, the respective signal is divided into a plurality of frequency band signals and thence fed to a band-dependent inter-channel time difference / level difference detector 5 and a sound source determination signal selector 6. Depending on each detection output from the detectors 3 and 5, the selector 6 selects a certain channel signal as A component or B component for each band. The selected A component signal and B component signal for each band are synthesized in sound source signal synthesizers 7A, 7B to be delivered separately as a sound source A signal and a sound source B signal.

When the sound source A is located closer to the microphone 1 than to the microphone 2, a signal SA1 from the source A reaches the microphone 1 earlier and at higher level than a signal SA2 from the sound source A reaches the microphone 2. Similarly, when the sound source B is located closer to the microphone 2 than to the microphone 1, a signal SB2 from the sound source B reaches the microphone 2 earlier, and at a higher level than a signal SB1 from the sound source B reaches the microphone 1. In this manner, in accordance with the invention, a variation in the acoustic signal reaching both microphones 1, 2 which is attributable to the locations of the sound sources relative to the microphones 1,2, or a difference in the time of arrival and a level difference between both signals, is utilized.

The operation of the apparatus as shown in Fig. 1 will be described with reference to Fig.2. As shown, signals from the two sound sources A, B are received by the microphones 1, 2 (S01). The inter-channel time difference / level difference detector 3 detects either an inter-channel time difference or a level difference from the L and R channel signals. As a parameter which is used in the detection of the time difference, the use of a cross-correlation function between the L and the R channel signal will be described below. Referring to Fig. 3, initially samples L(t) , R(t) of the L and the R signal are read (S02), and a cross-correlation function between these samples is calculated (S03). The calculation takes place by determining a cross-correlation at the same sampling point for the both channel signals, and then cross-correlations between the both channel signals when one of the channel signals is displaced by 1, 2 or more sampling points relative to the other channel signal. A number of such cross-correlations are obtained which are then normalized according to the power to form a histogram (S04). Time point differences $\Delta\alpha_1$ and $\Delta\alpha_2$ where the maximum and the second maximum in the cumulative frequency occur in the histogram are then determined (S05). These time point differences $\Delta\alpha_1$, $\Delta\alpha_2$ are then converted according to the equation given below into inter-channel time differences $\Delta\tau_1$, $\Delta\tau_2$ for delivery (S06).

$$\Delta\tau_1 = 1000 \times \Delta\alpha_1/F \quad (1)$$

$$\Delta\tau_2 = 1000 \times \Delta\alpha_2/F \quad (2)$$

where F represents a sampling frequency and a multiplication factor of 1000 is used to provide an increased magnitude for the convenience of calculation. The time differences $\Delta\tau_1$, $\Delta\tau_2$ represent inter-channel time differences in the L and R channel signal from the sound sources A, B.

Returning to Figs. 1 and 2, the bandsplitter 4 divides the L and the R signal into frequency band signals L(f1), L(f2), • • • , L(fn), and frequency band signals R(f1), R(f2), • • • , R(fn) (S04). This division may take place, for example, by using a discrete Fourier transform of each channel signal to convert it to a frequency domain signal, which is then divided into individual frequency bands. The bandsplitting takes place with a bandwidth, which may be 20 Hz, for example, for a voice signal, considering a difference in the frequency response of the signals from the sound sources A, B so that principally a signal component from only one sound source resides in each band. A power spectrum for the sound source A is obtained as illustrated in Fig. 4A, for example, while a power spectrum for the sound source B is obtained as illustrated in Fig. 4B. The bandsplitting takes place with a bandwidth $\Delta f$ of an order which permits the respective spectrums to be separated from each other. It will be seen then that as illustrated by broken lines connecting between corresponding spectrums, the spectrum for one of the sound sources is dominant, and the spectrum from the other sound source can be neglected. As will be understood from Figs. 4A and 4B, the bandsplitting may also take place with a bandwidth of $2\Delta f$. In other words, each band may not contain only one spectrum. It is also to be noted that the discrete Fourier transform takes place every 20 - 40 ms, for example.

The band-dependent inter-channel time difference / level difference detector 5 detects a band-dependent inter-channel time difference or level difference between the channels of each corresponding band signal such as L(f1) and R(f1), • • • L(fn) and R(fn), for example, (S05). The band-dependent inter-channel time difference is detected uniquely by utilizing the inter-channel time difference $\Delta\tau_1$, $\Delta\tau_2$ which are detected by the inter-channel time difference detector 3. This detection takes place utilizing the equations given below.



BNSDOCID: <EP_____0831458A2_I_>

$$\Delta\tau_1 - \{(\Delta\phi i/(2\pi fi)+(ki1/fi)\} = \varepsilon_i 1 \quad (3)$$

$$\Delta\tau_2 - \{(\Delta\phi i/(2\pi fi)+(ki2/fi)\} = \varepsilon_i 2 \quad (4)$$

where i = 1, 2, • • •, n, and $\Delta\phi i$ represents a phase difference between the signal L(fi) and the signal R(fi). Integers ki1, ki2 are determined so that $\varepsilon_i 1$, $\varepsilon_i 2$ assume their minimum values. The minimum values of $\varepsilon_i 1$ and $\varepsilon_i 2$ are compared against each other, and the smaller one of them is chosen as an inter-channel time difference $\Delta\tau_j$ (j = 1, 2), which represents an inter-channel time difference $\Delta\tau_{ij}$ for the band i. This represents an inter-channel time difference for one of the sound source signals in that band.

The sound source determination signal selector 6 utilizes the band-dependent inter-channel time differences $\Delta\tau_{1j}$ - $\Delta\tau_{nj}$ which are detected by the band-dependent inter-channel time difference / level difference detector 5 to render a determination in a sound source signal determination unit 601 which one of corresponding band signals L(f1) - L(fn) and R(f1) - R(fn) is to be selected ( S06 ). By way of example, an instance in which $\Delta\tau_1$ which is calculated by the inter-channel time difference / level difference detector 3 represents an inter-channel time difference for the signal from the sound source A which is located close to the microphone of the L side while $\Delta\tau_2$ represents an inter-channel time difference for the signal from the sound source B which is located close to the microphone for the R side will be described.

In this instance, for the band i for which the time difference $\Delta\tau_{ij}$ calculated by the band-dependent inter-channel time difference / level difference detector 5 is equal to $\tau_1$, the sound source signal determination unit 601 opens a gate 602 Li, whereby an input signal L(fi) of the L side is directly delivered as SA(fi) while for an input signal R(fi) for the band i of the R side, the sound source signal determination unit 601 closes a gate 602 R, whereby SB(fi) is delivered as 0. Conversely, for the band i for which the time difference $\Delta\tau_{ij}$ is equal to $\Delta\tau_2$, the signal L(fi) for the L side is delivered as SA(fi) = 0, and the input signal R(fi) for the R side is directly delivered as SB(fi). Thus, as shown in Fig. 1, the band signals L( f1) - L(fn) are fed to a sound source signal synthesizer 7A through gates 602L1 - 602Ln, respectively, while the band signal R(f1) - R(fn) are fed to a sound source signal synthesizer 7B through gates 602R1 - 602Rn, respectively. $\Delta\tau_{1j}$ - $\Delta\tau_{nj}$ are input to the sound source signal determination unit 601 within the sound source determination signal selector 6, and for the band i for which $\Delta\tau_{ij}$ is determined to be equal to $\Delta\tau_1$, gate control signals Cli = 1 and Cli = 0 are produced, thus controlling the corresponding gates 602Li and 602Ri to be opened and closed, respectively. For the band i for which $\Delta\tau_{ij}$ is determined to be equal to $\Delta\tau_2$, the gate control signals Cli = 0 and CRi = 1 are produced, controlling the corresponding gates 602Li and 602Ri to be closed and opened, respectively. It should be noted that the above description is given to describe the functional arrangement, but in practice, a digital signal processor, for example, is used to achieve the described operation.

The sound source signal synthesizer 7A synthesizes signals SA(fi) - SA(fn), which are subjected to an inverse Fourier transform in the above example of bandsplitting to be delivered to an output terminal $t_A$ as a signal SA. Similarly, the sound source signal synthesizer 7B synthesizes signals SB(fi) - SB(fn), which are delivered to an output terminal $t_B$ as a signal SB.

It will be apparent from the foregoing description that, in the apparatus of the invention, a determination is rendered as to from which sound source each band component which is finely divided from the respective channel signal accrues, and the components thus determined are all delivered. Thus, unless frequency components of signals from the sound sources A, B overlap each other, the processing operation takes place without dropping any specific frequency band, and accordingly, it is possible to separate the signals from the sound sources A, B from each other while maintaining a high voice quality as compared with a conventional process in which only harmonic structures are extracted.

In the foregoing description, the sound source signal determination unit 601 determined a condition for determination by merely utilizing an inter-channel time difference and a band-dependent inter-channel time difference which are detected by the inter-channel time difference / level difference detector 3 and the band-dependent inter-channel time difference / level difference detector 5.

Another embodiment in which the condition for determination is determined by using a inter-channel level difference will now be described. Such an embodiment is illustrated in Fig. 5. As shown, the L and the R channel signal are received by the microphones 1, 2, respectively ( S02 ), and inter-channel level difference $\Delta L$ between the L and the R channel signal is detected by the inter-channel time difference / level difference detector 3 ( Fig. 1) (S03). In a similar manner as occurs at the step S04 shown in Fig. 2, the L and the R channel signal are each divided into n band-dependent channel signals L(f1) - L(fn) and R(f1) - R(fn) (S04), and band-dependent inter-channel level differences $\Delta L1$, $\Delta L2$, • • •, $\Delta Ln$ between corresponding bands in the band-dependent channel signals L(f1) - L(fn) and R(f1) - R(fn) or between L(f1) and R(f1), between L(f2) and R(f2), • • • and between L(fn) and R(fn) are detected (S05).

A human voice can be considered to remain in its steady state condition during an interval on the order 20 - 40 ms. Accordingly, the sound source signal determination unit 601 ( Fig.1 ) calculates, every interval of 20 - 40 ms, the percentage of bands relative to all the bands in which the sign of the logarithm of the inter-channel level difference $\Delta L$ and the sign of the logarithm of the band-dependent inter-channel level difference $\Delta Li$ is equal ( either + or - ). If the percent-



BNSDOCID: <EP_____0831458A2_I_>

age is above a given value, for example, equal to or greater than 80 % ( S06, S07), the determination takes place only according to the inter-channel level difference $\Delta L$ for a subsequent interval of 20 - 40 ms( S08 ). If the percentage is less than 80 %, the determination takes place according the band-dependent inter-channel level difference $\Delta Li$ for every band during a subsequent interval of 20 - 40 ms (S09). The determination takes place in a manner such that when the determination takes place according to the inter-channel level difference $\Delta L$ for all the bands and when $\Delta L$ is positive, the L channel signal L(t) is directly delivered as the signal SA while the R channel signal R(t) is delivered as a signal SB = 0. Conversely, if $\Delta L$ is equal to or less than 0, the L channel signal L(t) is delivered as the signal SA = 0 while the R channel signal R(t) is directly delivered as the signal SB. However, it should be understood that this applies when a value which is obtained by subtracting the R side from the L side is used as the inter-channel level difference. When the determination takes place for each band using the band-dependent inter-channel level difference $\Delta Li$, the L side divided signal L(fi) are directly delivered as the signal SA(fi) while the R side divided signals R(fi) are delivered as signal SB(fi) equal to 0 when the band-dependent inter-channel level difference $\Delta Li$ for each band fi is positive. When the level difference $\Delta Li$ is equal to or less than 0, the L side divided signals L(fi) are delivered as signal SA(fi) equal to 0 while the R side divided signals R(fi) are delivered as signal SB(fi ). In this manner, the sound source signal determination unit 601 provide gate control signals CL1 - CLn, CR1 - CRn, which control gates 602 L1-602 Ln, 602 R1 - 602 Rn, respectively. As mentioned previously, this description applies when a value obtained by subtracting the R side from the L side is used for the band-dependent inter-channel level difference. As in the previous embodiment, the signals SA(f1) - SA(fn) and signals SB(f1) - SB(fn) are delivered to output terminals $t_A$, $t_B$, respectively, as synthesized signals SA, SB ( S10 ).

In the above embodiment, only one of a difference in the time of arrival and the level difference is utilized as the condition for determination which is used in the sound source signal determination unit 601. However, when only the level difference is used, it is possible that the levels of L(fi) and R(fi) compare equally in low frequency bands, and it is then difficult to determine the level difference accurately. Also, when only the time difference is used, a phase rotation presents a difficulty in correctly calculating the time difference in high frequency bands. In view of these, it may be advantageous to use the time difference in low frequency bands and to use the level difference in high frequency bands for the determination rather than using a single parameter over the entire band.

Accordingly, a further embodiment in which the band-dependent inter-channel time difference and band-dependent inter-channel level difference are both used in the sound source signal determination unit 601 will be described with reference to Fig. 6 and subsequent Figures. A functional block diagram for this arrangement remains the same as shown in Fig. 1, but a processing operation which takes place in the inter-channel time difference / level difference detector 3, the band-dependent inter-channel time difference / level difference detector 5 and the sound source signal determination unit 601 becomes different as mentioned below. The inter-channel time difference / level difference detector 3 delivers a single time difference $\Delta\tau$ such as a mean value of absolute magnitudes of the detected time differences $\Delta\tau_1$, $\Delta\tau_2$ or only one of $\Delta\tau_1$, $\Delta\tau_2$ if they are relatively close to each other. It is to be noted that while the inter-channel time differences $\Delta\tau_1$, $\Delta\tau_2$, $\Delta\tau$ are calculated before the channel signals L(t), R(t) are devided into bands on the frequency axis, it is also possible to calculate such time differences after the bandsplitting.

Referring to Fig. 5, the L channel signal L(t) and the R channel signal R(t) are read every frame ( which may be 20 - 40 ms, for example ) ( S02 ), and the bandsplitter 4 divides the L and R channel signals into a plurality of frequency bands, respectively. In the present example, a Humming window is applied to the L channel signal L(t) and the R channel signal R(t) (S03), and then they are subject to a Fourier transform to obtain divided signals L(f1) - L(fn), R(f1) - R(fn) (S04).

The band-dependent inter-channel time difference / level difference detector 5 then examines if the frequency fi of the divided signal is a band ( hereafter referred to as a low band ) which corresponds to $1/(2\Delta\tau)$ ( where $\Delta\tau$ represents a channel time difference ) or less ( S05 ). If this is the case, a band-dependent inter-channel phase difference $\Delta\phi i$ is delivered (S08). It is then examined if the frequency f of the divided signal is higher than $1/(2\Delta\tau)$ and less than $1/\Delta\tau$ ( hereafter referred to as a middle band ) ( S06 ). If the frequency lies in the middle band, the band-dependent interchannel phase difference $\Delta\phi i$ and level difference $\Delta Li$ are delivered ( S09 ). Finally, it is examined if the frequency f of the divided signal lies in a band corresponding to $1/\Delta\tau$ or higher ( hereafter referred to as a high band ) ( S07 ), and for the high band, the band-dependent inter-channel level difference $\Delta Li$ is delivered ( S10 ).

The sound source signal determination unit 601 uses the band-dependent inter-channel phase difference and the level difference which are detected by the band-dependent inter-channel time difference / level difference detector 5 to determine which one of L(f1) - L(fn) and R(f1) - R(fn) is to be delivered. It is to be noted that a value which is obtained by subtracting the R side value from the L side value is used for the phase difference $\Delta\phi i$ and the level difference $\Delta L$ in the present example.

Referring to Fig. 7, for signals L(fi), R(fi) which are determined as lying in the low band, an examination is initially made to see if the phase difference $\Delta\phi i$ is equal to or greater than $\pi$ ( S15 ). If the phase difference is equal to or greater than $\pi$, $2\pi$ is subtracted from $\Delta\phi i$ to update $\Delta\phi i$ ( S17 ). If it is found at step S15 that $\Delta\phi i$ is less than $\pi$, an examination is made to see if it is equal to or less than $-\pi$ (S16). If it is equal to or less than $-\pi$, $2\pi$ is added to $\Delta\phi i$ to update $\Delta\phi i$ (



BNSDOCID: <EP_____0831458A2_I_>

S18 ). If it is found at step S16 that the phase difference is not equal to or less than - $\pi$, $\Delta\phi i$ is used without change ( S19 ). The band-dependent inter-channel phase difference $\Delta\phi i$ which is determined at steps S17, S18 and S19 is converted into a time difference $\Delta\sigma i$ according to the equation given below ( S20 ).

$$\Delta\phi i = 1000 \times \Delta\phi i/2\pi fi \quad (5)$$

When the divided signals L(fi) , R(fi) are determined as lying in the middle band, the phase difference $\Delta\phi i$ is determined uniquely by utilizing the band-dependent inter-channel level difference $\Delta L(fi)$ as indicated in Fig.8. Specifically, an examination is made to see if $\Delta L(fi)$ is positive ( S23 ), and if it is positive, an examination is again made to see if the band-dependent inter-channel phase difference $\Delta\phi i$ is positive ( S24). If the phase difference is positive, this $\Delta\phi i$ is directly delivered ( S26 ). If it is found at step S24 that the phase difference is not positive, $2\pi$ is added to $\Delta\phi i$ to update it ( S27 ). If it is found at step S23 that $\Delta L(fi)$ is not positive, an examination is made to see if the band-dependent inter-channel phase difference $\Delta\phi i$ is negative ( S25 ), and if it is negative, this $\Delta\phi i$ is directly delivered ( S28 ). If it is found at step S25 that the phase difference is not negative, $2\pi$ is subtracted from $\Delta\phi i$ to update it for delivery ( S29 ). $\Delta\phi i$ which is determined at one of the steps S26 to S29 is used in the equation given below to determine a band-dependent inter-channel time difference $\Delta\sigma i$ ( S30 ).

$$\Delta\phi i = 1000 \times \Delta\phi i/2\pi fi \quad (6)$$

In the manner mentioned above, the band-dependent inter-channel time difference $\Delta\phi i$ in the low and the middle band as well as the band-dependent inter-channel level difference $\Delta L(fi)$ in the high band are obtained, and sound source signal is determined in accordance with these variables in a manner mentioned below.

Referring to Fig. 9, by utilizing the phase difference $\Delta\phi i$ in the low and the middle band and utilizing the level difference $\Delta Li$ in the high band, the respective frequency components of both channels are determined as signals of either applicable sound source, in a manner shown in Fig.9. Specifically, for the low and the middle band, an examination is made to see if the band-dependent inter-channel time difference $\Delta\phi i$ which is determined in manners illustrated in Figs. 7 and 8 is positive ( S34 ), and if it is positive, the L side channel signal L(fi) of the band i is delivered as the signal SA(fi) while the R side band channel signal R(fi) is delivered as the signal SB(fi) of 0 ( S36 ). Conversely, if it is found at step S34 that band-dependent inter-channel time difference $\Delta\phi i$ is not positive, SA(fi) is delivered as 0 while the R side channel signal R(fi) is delivered as SB(fi) ( S37 ).

For the high band, an examination is made to see if the band-dependent inter-channel level difference $\Delta L(fi)$ which is detected at step S10 in Fig. 6 is positive ( S35 ), and if it is positive, the L side channel signal L(fi) is delivered as signal SA(fi) while 0 is delivered as SB(fi) ( S38 ). If it is found at step S35 that the level difference $\Delta Li$ is not positive, 0 is delivered as signal SA(fi) while the R side channel signal R(fi) is delivered as SB(fi) ( S39 ).

In the manner mentioned above, the L side or R side signal is delivered from the respective bands, and the sound source signal synthesizers 7A, 7B add the frequency components thus determined over the entire band ( S40 ) and the added sum is subjected to the inverse Fourier transform ( S41 ), thus delivering the transformed signals SA, SB ( S42 ).

In the present embodiment, by utilizing a parameter which is preferred for the separation of the sound source for every frequency band in the manner mentioned above, it is possible to achieve the separation of a sound source with a higher separation performance than when a single parameter is used over the entire band.

The invention is also applicable to three or more sound sources. By way of example, the separation of sound source when the number of sound sources is equal to three and the number of microphones is equal to two by utilizing the difference in the time of arrival to the microphones will be described. In this instance, when the inter-channel time difference / level difference detector 3 calculates an inter-channel time difference for the L and the R channel signal for each sound source, the inter-channel time differences $\Delta\tau_1$, $\Delta\tau_2$, $\Delta\tau_3$ for the respective sound source signals are calculated by determining points in time when a first rank to a third rank peak in the cumulative frequency occurs in the histogram which is normalized by the power of the cross-correlations as illustrated in Fig. 3. Also, the band-dependent inter-channel time difference / level difference detector 5 determines the band-dependent inter-channel time difference for each band as to be one of $\Delta\tau_1$ to $\Delta\tau_3$. This manner of determination remains similar as used in the previous embodiments using the equations (3), (4). The operation of the sound source signal determination unit 601 will be described for an example in which $\Delta\tau_1>0$, $\Delta\tau_2>0$, $\Delta\tau_3<0$. It is assumed that $\Delta\tau_1$, $\Delta\tau_2$, $\Delta\tau_3$ represent the inter-channel time differences for the signals from the sound sources A, B, C, respectively, and it is also assumed that these values are derived by subtracting the R side value from the L side value. In this instance, the sound source A is located close to the L side microphone 1 while the sound source B is located close to the R side microphone 2. Thus, it is possible to separate the signal from the sound source A on the basis of the L channel signal, to which a signal for the band where the band-dependent inter-channel time difference is equal to $\Delta\tau_1$ is added, and to separate the signal for the sound source B on the basis of the L channel signal, to which the signal for the band in which the band-dependent inter-channel time difference is equal to $\Delta\tau_2$ is added. The signal from the sound source C is separated on the basis of the R channel signal, to which the



BNSDOCID: <EP_____0831458A2_I_>

signal for the band in which the band-dependent inter-channel time difference is equal to $\Delta\tau_3$ is added.

In the above description, sound source signals are separated, and the separated sound source signals SA, SB have been separately delivered. However, if one of the sound sources, A, is a voice uttered by a speaker while the other sound source B represents a noise, the invention can be applied to separate and extract the signal from the sound source A from the mixture with the noise while suppressing the noise. In such an instance, the sound source signal synthesizer 7A may be left while the sound source signal synthesizer 7B, gates 602R1 - 602Rn shown within a dotted line frame 9 may be omitted in the arrangement of Fig. 1.

Where the frequency band of one of the sound sources, A, is broader than the frequency band of the other sound source B and the respective frequency bands are previously known, a band separator 10 as shown in Fig. 10 may be used in the arrangement of Fig. 1 to separate a frequency band where there is no overlap between both sound source signals. To give an example, it is assumed that the signal A(t) of the sound source A has a frequency band of f1 - fn while the signal B(t) from the sound source B has a frequency band of f1 - fn (where fn > fm). In this instance, a signal in the non-overlapping band fm + 1 - fn can be separated from the outputs of the microphones 1, 2. The sound source signal determination unit 601 does not render a determination as to the signal in the band fm + 1 - fn, and optionally a processing operation by the band-dependent inter-channel time difference / level difference detector 5 may also be omitted. The sound source signal determination unit 601 controls the sound source signal selector 602 in a manner such that the R side divided band channel signals R(fm + 1) - R(fn), which are selected as channel signal SB(t) from the sound source B, are delivered as SB(fm + 1) - SB(fn) while 0 is delivered as SA(fm + 1) - SA(fn). Thus, gates 602Lm + 1 - 602Ln are normally closed while gates 602Rm + 1 - 602Rn are normally open.

In the foregoing description, a determination has been rendered to which microphone a particular band signal is close depending on the positive or negative polarity of the respective band-dependent inter-channel time difference $\Delta\sigma i$ or the positive or negative polarity of the respective band-dependent inter-channel level difference $\Delta Li$, thus using 0 as a threshold. This applies when the sound sources A and B are symmetrically located on the opposite sides of a bisector of a line joining the microphone 1. Where this relationship does not apply, a threshold can be determined in a manner mentioned below.

A band-dependent inter-channel level difference and band-dependent inter-channel time difference when a signal from the sound source A reaches the microphones 1 and 2 are denoted by $\Delta L_A$ and $\Delta\tau_A$ while a band-dependent inter-channel level difference and band-dependent inter-channel time difference when a signal from the sound source B reaches the microphones 1 and 2 are denoted by $\Delta L_B$ and $\Delta\tau_B$, respectively. At this time, a threshold $\Delta$Lth for the band-dependent inter-channel level difference may be chosen as

$$\Delta Lth = (\Delta L_A + \Delta L_I)/2$$

and a threshold value $\Delta\tau$th for the band-dependent inter-channel time difference may be chosen as

$$\Delta\tau th = (\Delta\tau_A + \Delta\tau_B)/2$$

In the embodiment mentioned previously, $\Delta L_B = -\Delta L_A$, $\Delta\tau_B = -\Delta\tau_A$. Hence, $\Delta$Lth = 0 and $\Delta\tau$th = 0. The microphones 1, 2 are located so that the two sound sources are located on opposite sides of the microphones 1,2 in order that a good separation between the sound sources can be achieved. However, under certain circumstances, the distance and direction with respect to the microphones 1, 2 can not be accurately known and in such instance, the thresholds $\Delta$Lth, $\Delta\tau$th may be chosen to be variable so that these thresholds are adjustable to enable a good separation.

It is possible with the described embodiments that an error may occur in the band-dependent inter-channel time difference or band-dependent inter-channel level difference under the influence of reverberations or diffractions occurring in the room, preventing a separation of the respective sound source signals from being achieved with a good accuracy. Another embodiment which accommodates for such a problem will now be described. In an example shown in Fig. 11, microphones M1, M2, M3 are disposed at the apices of an equilateral triangle measuring 20 cm on a side, for example. The space is divided in accordance with the directivity of the microphones M1 to M3, and each divided sub-space is referred to as a sound source zone. Where all of the microphones M1 to M3 are non-directional and exhibit similar response, the space is divided into six zones Z1 - Z6, as illustrated in Fig. 12, for example. Specifically, six zones Z1 - Z6 are formed about a center point Cp at an equi-angular interval by rectilinear lines, each passing the respective microphones M1, M2, M3 and the center point Cp. The sound source A is located within the zone Z3 while the sound source B is located within the zone Z4. In this manner, the individual sound source zones are determined on the basis of the disposition and the responses of the microphones M1 - M3 so that one sound source belongs to one sound source zone.

Referring to Fig. 11, a bandsplitter 41 divides an acoustic signal S1 of a first channel which is received by the microphone M1 into n frequency band signals S1(f1) - S1(fn). A bandsplitter 42 divides an acoustic signal S2 of a second channel which is received by the microphone M2 into n frequency band signals S2(f1) - S2(fn), and a bandsplitter 43



BNSDOCID: <EP_____0831458A2_I_>

divides an acoustic signal S3 of a third channel which is received by the microphone M3 into n frequency band signals S3(f1) - S3(fn). The bands f1 - fn are common to the bandsplitters 41 - 43 and a discrete Fourier transform may be utilized in providing such bandsplitting.

A sound source separator 80 separates a sound source signal using the techniques mentioned above with reference to Figs. 1 to 10. It should be noted, however, that since there are three microphones in the arrangement of Fig. 11, a similar processing as mentioned above is applied to each combination of two of the three channel signals. Accordingly, the bandsplitters 41 - 43 may also serve as bandsplitters within the sound source separator 80.

A band-dependent level ( power ) detector 51 detects level ( power ) signals P( S1f1) - P( S1fn ) for the respective band signals S1(f1) - S1(fn) which are obtained by the bandsplitter 41. Similarly, band-dependent level detectors 52, 53 detect the level signals P(S2f1) - P(S2fn), P(S3f1) - P(S3fn) for the band signals S2(f1) - S2(fn), S3(f1) - S3(fn) which are obtained in the bandsplitters 42, 43, respectively. The band-dependent level detection can also be achieved by using the Fourier transforms. Specifically, each channel signal is resolved into a spectrum by the discrete Fourier transform, and the power of the spectrum may be determined. Accordingly, a power spectrum is obtained for each channel signal, and the power spectrum may be band splitted. The channel signals from the respective microphones M1 - M3 may be band splitted in a band-dependent level detector 400, which delivers the level ( power ).

On the other hand, an all band level detector 61 detects the level (power)P(S1) of all the frequency components contained in an acoustic signal S1 of a first channel which is received by the microphone M1. Similarly, all band level detectors 62, 63 detect levels P(S2), P(S3) of all frequency components of acoustic signals S2, S3 of second and third channels 2, 3 which are received by the microphones M2, M3, respectively.

A sound source status determination unit 70 determines, by a computer operation, any sound source zone which is not uttering any acoustic sound. Initially, the band-dependent levels P(S1f1) - P(S1fn), P(S2f1) - P(S2fn) and P(S3f1) - P(S3fn) which are obtained by the band-dependent level detector 50 are compared against each other for the same band signals. In this manner, a channel which exhibits a maximum level is specified for each band f1 to fn.

By choosing a number n of the divided bands which is above a given value, it is possible to choose an arrangement in which a single band only contains an acoustic signal from single sound source as mentioned previously, and accordingly, the levels P(S1fi), P(S2fi), P(S3fi) for the same band fi can be regarded as representing acoustic levels from the same sound source. Consequently, whenever there is a difference between the P(S1fi), P(S2fi), P(S3fi) for the same band between the first to the third channel, it will be seen that the level for the band which comes from a microphone channel located closest to the sound source is at maximum.

As a result of the preceding processings, a channel which exhibits the maximum level is allotted to each of the bands f1 - fn. A total number of bands $\chi 1$, $\chi 2$, $\chi 3$ for which each of the first to the third channel exhibited the maximum level among n bands is calculated. It will be seen that the microphone of the channel which has a greater total number is located close to the sound source. If the total number is on the order of 90n/100 or greater, for example, it may be determined that the sound source is close to the microphone of that channel. However, if a maximum total number of highest level bands is equal to 53n/100, and a second maximum total number is equal to 49n/100, it is not certain if the sound source is located close to a corresponding microphone. Accordingly, a determination is rendered such that the sound source is located closest to the microphone of a channel which corresponds to the total number when the total number is at maximum and exceeds a preset reference value ThP, which may be on the order of n/3, for example.

The levels P(S1) - P(S3) of the respective channels which are detected by the all band level detector 60 is also input to the sound source determination unit 70, and when all the levels are equal to or less than a preset value ThR, it is determined that there is no sound source in any zone.

On the basis of a result of determination rendered by the sound source status determination unit 70, a control signal is generated to effect a suppression upon acoustic signals A, B which are separated by the sound source separator 80 in a signal suppression unit 90. Specifically, a control signal SAi is used to suppress ( attenuate or eliminate ) an acoustic signal SA; a control signal SBi is used to suppress an acoustic signal SB; and a control signal SABi is used to suppress both acoustic signals SA, SB. By way of example, the signal suppression unit 90 may include normally closed switches 9A, 9B, through which output terminals $t_A$, $t_B$ of the sound source separator 80 are connected to output terminals $t_{A'}$, $t_{B'}$. The switch 9A is opened by the control signal SAi, the switch 9B is opened by the control signal SBi, and both switches 9A, 9B are opened by the control signal SABi. Obviously, the frame signal which is separated in the sound source separator 80 must be the same as the frame signal from which the control signal used for suppression in the signal suppression unit 90 is obtained. The generation of suppression ( control ) signals SAi, SBi, SABi will be described more specifically.

When the sound sources A, B are located as shown in Fig. 12, microphones M1 - M3 are disposed as illustrated to determine zones Z1 - Z6 so that the sound sources A and B are disposed within separate zones Z3 and Z4. It will be seen that at this time, the distances SA1, SA2, SA3 from the sound source A to the microphones M1 - M3 are related such that SA2 < SA3 < SA1. Similarly, distances SB1, SB2, SB3 from the sound source B to the respective microphones M1 - M3 are related such that SB3 < SB2 < SB1.

When all of the detection signals P(S1) - P(S3) from the all band level detector 60 are less than the reference value



BNSDOCID: <EP_____0831458A2_I_>

ThR, the sound sources A, B are regarded as not uttering a voice or speaking, and accordingly, the control signal SABi is used to suppress both acoustic signals SA, SB. At this time, the output acoustic signals SA, SB are silent signals (see blocks 101 and 102 in Fig. 13).

When only the sound source A is uttering a voice, its acoustic signal reaches the microphone M2 at a maximum sound pressure level (power) for the frequency component of all the bands, and accordingly, the total number of bands $\chi 2$ for the channel corresponding to the microphone M2 is at maximum.

When only the sound source B is uttering a voice, its acoustic signal reaches the microphone M3 at a maximum sound pressure level for frequency components of all the bands, and accordingly the total number of bands $\chi 3$ for the channel corresponding the microphone M3 is at maximum.

When both sound sources A, B are uttering a voice, the number of bands in which the acoustic signal reaches the maximum sound pressure level will be comparable between the microphones M2 and M3.

Accordingly, when the total number of bands in which the acoustic signal reaches the microphone at the maximum sound pressure level exceeds the reference value ThP mentioned above, a determination is rendered that there exists a sound source in the zone which is covered by this microphone, thus enabling a sound source zone in which an utterance of a voice is occurring to be detected.

In the above example, if only the sound source A is uttering a voice, only $\chi 2$ will exceed the reference value ThP, thus providing a detection that the uttering sound source exists only in the zone Z3 covered by the microphone M2. Accordingly, the control signal SBi is used to suppress the voice signal SB while allowing only the acoustic signal SA to be delivered (see blocks 103 and 104 in Fig.13).

Where only the sound source B is uttering a voice, $\chi 3$ will exceed the reference value ThP, providing a detection that the uttering sound source exists in the zone Z4 covered by the microphone M3, and accordingly, the control signal SAi is used to suppress the acoustic signal SA while allowing the acoustic signal SB to be delivered alone (see blocks 105 and 106 in Fig. 13).

Finally, when both the sound sources A, B are uttering a voice, and when both $\chi 2$ and $\chi 3$ exceed the reference value ThP, a preference may be given to the sound source A, for example, treating this case as the utterance occurring only from the sound source A. The processing procedure shown in Fig. 13 is arranged in this manner. If both $\chi 2$ and $\chi 3$ fail to reach the reference value ThP, it may be determined that both sound sources A, B are uttering a voice as long as the levels P(S1) - P(S3) exceed the reference value ThR. In this instance, none of the control signals SAi, SBi, SABi is delivered, and the suppression of the synthesize signals SA, SB in the signal suppression unit 90 does not take place (see block 107 in Fig. 13).

In this manner, the sound source signals SA, SB which are separated in the sound source separator 80 are fed to the sound source status determination unit 70 which may determine that a sound source is not uttering a voice, and a corresponding signal is suppressed in the signal suppression unit 90, thus suppressing unnecessary sound.

A sound source C may be added to the zone Z6 in the arrangement shown in Fig. 12, as illustrated in Fig. 14. While not shown, in this instance, the sound source separator 80 delivers a signal SC corresponding to the sound source C in addition to the signals SA, SB,corresponding the sound sources A, B, respectively.

The sound source status determination unit 70 delivers a control signal SCi which suppresses the signal SC to the signal suppression unit 90, in addition the control signal SAi which suppresses the signal SA and the control signal SBi which suppresses the signal SB. Also, in addition to the control signal SABi which suppresses both the signal SA and the signal SB, a control signal SBCi which suppresses the signals SB, SC, a control signal SCAi which suppresses the signals SC, SA, and a control signal SABCi which suppresses all of the signals SA, SB, SC are delivered. The sound source status determination unit 70 operates in a manner illustrated in Fig. 15.

Initially, if none of the levels P(S1) - P(S3) exceed the reference ThR, a determination is rendered that none of the sound sources A to C are uttering a voice, and accordingly the sound source status determination unit 70 delivers the control signal SABCi, suppressing all of the signals SA, SB, SC (see blocks 201 and 202 in Fig. 15).

Then, if the sound source A, B or C is uttering a voice alone, one of the levels P(S1) - P(S3) exceeds the reference value ThR, and the level of the channel corresponding to the microphone which is located closest to the uttering sound source will be at maximum, in a similar manner as when there are two sound sources mentioned above, and accordingly, one of the channel band number $\chi 1$, $\chi 2$, $\chi 3$ will exceed the reference value ThP. If only the sound source C is uttering a voice, $\chi 1$ will exceed ThP, whereby the control signal SABi is delivered to suppress the signals SA, SB (see blocks 203 and 204 in Fig.15). If only the sound source A is uttering a voice, the control signal SBCi is delivered to suppress the signals SB, SC. Finally, if only the sound source B is uttering a voice the control signal SACi is delivered to suppress the signals SA, SC (see blocks 205 to 208 in Fig. 15).

When any two of the three sound sources A to C are uttering a voice, the total number of bands in which the channel corresponding to the microphone located in a zone corresponding to the non-uttering sound source exhibits a maximum level will be reduced as compared with the other microphones. For example, when only the sound source C is not uttering a voice, the total number of bands $\chi 1$ in which the channel corresponding to the microphone M1 exhibits the maximum level will be reduced as compared with the total number of bands $\chi 2$, $\chi 3$ corresponding to other microphones



BNSDOCID: <EP_____0831458A2_I_>

M2, M3.

In consideration of this, a reference value ThQ (<ThP) may be established, and if $\chi 1$ is equal to or less than the reference value ThQ, a determination is rendered that of the zones Z5, Z6 each of which is bisected by the microphone M1 and M3, respectively, a sound source is not producing a signal in the zone Z6 which is located close to the microphone M1. In addition, of the zones Z1, Z2 which are bisected by the microphone M1 and M2, respectively, a determination is rendered that in zone Z1 located close to the microphone M1, sound source is not producing a signal.

In this manner, a sound source located in the zones Z1, Z6 is determined as not producing a signal. Since the sound source located in such zones represents the sound source C, it is determined that the sound source C is not producing a signal or that only the sound sources A, B are producing a signal. Accordingly, the control signal SCi is generated, suppressing the signal SC. In the arrangement shown in Fig. 14, if only one of the three sound sources A to C fail to utter a voice, the total number of bands $\chi 1$, $\chi 2$, $\chi 3$ which either microphone exhibits a maximum level will normally be equal to or less than the reference value ThP. Accordingly, steps 203, 205 and 207 shown in Fig. 15 are passed, and an examination is made at step 209 if $\chi 1$ is equal to or less than the reference value ThQ. If it is found that only the sound source C does not utter a voice, it follows $\chi 1 <$ ThQ, generating the control signal SCi (see 210 in Fig. 15). If it is found at step 209 that $\chi 1$ is not less than ThQ, a similar examination is made to see if $\chi 2$ , $\chi 3$ is equal to or less than ThQ. If either one of them is equal to or less than ThQ, it is estimated that only the sound source A or only the sound source B fail to utter a voice, thus generating the control signal SAi or SBi (see 211 to 214 in Fig. 15).

When it is determined at step 213 that $\chi 3$ is not less than ThQ, a determination is rendered that all of the sound sources A, B, C are uttering a voice, generating no control signal (see 215 in Fig. 15).

In this instance, assuming that ThP is on the order of 2n/3 to 3n/4, the reference value ThQ will be on the order of n/2 to 2n/3, or if ThP is on the order of 2n/3, ThQ will be on the order of n/2.

In the above example, the space is divided into six zones Z1 to Z6. However, the status of the sound source can be similarly determined if the space is divided into three zones Z1 - Z3 as illustrated by dotted lines in Fig. 16 which pass through the center point Cp and through the center of the respective microphones. In this instance, if only the sound source A is uttering a voice, for example, the total number of bands $\chi 2$ of the channel corresponding to the microphone M2 will at maximum, and a determination is rendered that there is a sound source in the zone Z2 covered by the microphone M2. When only the sound source B is uttering a voice, $\chi 3$ will be at maximum, and a determination is rendered that there is a sound source in the zone Z3. If $\chi 1$ is equal to or less than the preset value ThQ, a determination is rendered that a sound source located in the zone Z1 is not uttering a voice. By the operation mentioned above, when the space is divided into three zones, the status of a sound source can be determined in similar manner as when the space is divided into six zones.

In the above description, the reference values ThR, ThP, ThQ are used in common for all of the microphones M1 - M3, but they may be suitably changed for each microphone. In addition, while in the above description, the number of sound sources is equal to three and the number of microphones is equal to three, a similar detection is possible if the number of microphones is equal to or greater than the number of sound sources.

For example, when there are four sound sources, the space is divided into four zones in a similar manner as illustrated in Fig.16 so that the four microphones may be used in a manner such that the microphone of each individual channel covers a single sound source. The determination of the status of the sound source in this instance takes place in a similar manner as illustrated by steps 201 to 208 in Fig. 15, thus determining if all of the four sound sources are silent or if one of them is uttering a voice. Otherwise, a processing operation takes place in a similar manner as illustrated by steps 209 to 214 shown in Fig. 15, determining if one of the four sound sources is silent, and in the absence of any silent sound source, a processing operation similar to that illustrated by the step 215 shown in Fig. 15 is employed, rendering a determination that all of the sound sources are uttering a voice.

Where three of the four sound sources are uttering a voice (or when one of the sound sources remains silent), no additional processing can be dispensed with, however, to discriminate one of the three sound sources which is more close to the silent condition, a fine control may take place as indicated below. Specifically, the reference value is changed from ThQ to ThS (ThP > ThS > ThQ) and each of the steps 210, 212, 214 shown in Fig. 15 may be followed by a processor as illustrated by steps 209 to 214 shown in Fig. 15, thus determining one of the three sound sources which is more close to the silent condition.

In this manner, as the number of sound sources increases, the processing operation illustrated by the steps 209 to 214 shown in Fig. 15 may be repeated to determine two or more sound sources which remain silent or which are close to a silent condition. However, as the number of repetitions increases, the reference value ThS used in the determination is made closer to ThP.

The procedure of processing operation for the described arrangement will be as shown in Fig. 17 when there are four microphones and four sound sources. Initially, a first to a fourth channel signal S1 - S4 are received by microphones M1 - M4 (S01), the levels P(S1) - P(S4)of theses channel signals S1 - S4 are detected (S02), an examination is made to see if these levels P(S1) - P(S4) are equal to or less than the threshold value ThR (S03), and if they are equal to or less than the reference value, a control signal SABCDi is generated to suppress synthesized signals SA, SB, SC (S1)



BNSDOCID: <EP_____0831458A2_I_>

from being delivered (S04). If it is found at step S03 that either one of the levels P(S1) - P(S4) is not less than the reference value ThR, the respective channel signal S1 - S4 are divided in to n bands, and the levels P(S1fi), P(S2fi), P(S3fi), P(S4fi), where (i = 1, • • •, n) of the respective bands are determined (S05). For each band fi, a channel fiM (where M is one of 1, 2, 3 or 4) which exhibits a maximum level is determined (S06), and the total number of bands for fi1, fi2, fi3, fi4, which are denoted as $\chi 1$, $\chi 2$, $\chi 3$, $\chi 4$, are determined among n bands (S07). A maximum one $\chi_M$ among $\chi 1$, $\chi 2$, $\chi 3$, and $\chi 4$ is determined (S08), an examination is made to see if $\chi_M$ is equal to or greater than the reference value ThP1 (which may be equal to n/3, for example) (S09), and if it is equal to or greater than ThP1, the sound source signal which is selected in correspondence to the channel M is delivered while generating a control signal SBCDi assuming that the sound source corresponding to channel M is sound source A which suppresses acoustic signals of separated channels other than channel M (S010). The operation may directly transfer from step S08 to step S010.

If it is found at step S09 that $\chi_M$ is not equal to or greater than the reference value, an examination is made to see if there is a channel M having $\chi_M$ which is equal to or less than the reference value ThQ (S011). If there is no such channel, all the sound sources are regarded as uttering a voice, and hence no control signal is generated (S012). If it is found at step S011 that there is a channel M having $\chi_M$ which is equal to or less than ThQ, a control signal SMi which suppress the sound source which is separated as the corresponding channel M is generated (S013).

There may be the separated sound source signal or signals other than the one suppressed by the control signal SMi which remains silent or which remains close to a silent condition. In order to suppress such sound source signal or signals, S is incremented by 1 (S014) (It being understood that S is previously initialized to 0), an examination is made to see if S matches M minus 1 (where M represents the number of sound sources) (S015), and if it does not match, ThQ is increased by an increment $+\Delta Q$ and the operation returns to step S011 (S016). The step S011 is repeatedly executed while increasing ThQ by an increment of $\Delta Q$ within the constraint that it does not exceed ThP until S becomes equal to M minus 1. If it is found at step S015 that M minus 1 equals S, each control signal SMi which suppresses a separated sound source signal corresponding to each channel for which $\chi_M$ is equal to or less than ThQ is generated (S013). If necessary, the operation may transfer to step S013 before M - 1 = S is reached at step S015.

After calculating $\chi 1$ - $\chi 4$ at step S07, an examination is made to see if there is any one which is above ThP2 (which may be equal to2n/3, for example). If there is such a one, the operation transfers to step S010, and otherwise the operation may proceed to step S011 (S016).

In the foregoing description, a control signal or signals for the signal suppression unit 90 is generated utilizing the inter-band level differences of the channels S1 - S3 corresponding to the microphones M1 - M3 in order to enhance the accuracy of separating the sound source. However, it is also possible to generate a control signal by utilizing an inter-band time difference.

Such an example is shown in Fig. 18 where corresponding parts to those shown in Fig. 11 are designated by like reference numerals and characters as used before. In this embodiment, a time-of-arrival difference signal An(S1f1) - An(S1fn) is detected by a band-dependent time difference detector 101 from signals S1(f1) - S1(fn) for the respective bands f1 - fn which are obtained in the bandsplitter 41. Similarly, time-of-arrival difference signals An(S2f1) - An(S2fn), An(S3f1) - An(S3fn) are detected by the band-dependent time difference detectors 102, 103, respectively, from the signals S2(f1) - S2(fn), S3(f1) - S3(fn) for the respective bands which are obtained in the bandsplitters 42, 43, respectively.

The procedure for obtaining such a time-of-arrival difference signal may utilize the Fourier transform, for example, to calculate the phase (or group delay) of the signal of each band followed by a comparison of the phases of the signals S1(fi), S2(fi), S3(fi) (where i equals 1, 2, • • •, n) for the common band fi against each other to derive a signal which corresponds to a time-of-arrival difference for the same sound source signal. Here again, the bandsplitter 40 uses a subdivision which is small enough to assure that there is only one sound source signal component in one band.

To express such a time-of-arrival difference, one of the microphones M1 - M3 may be chosen as a reference, for example, thus establishing a time-of-arrival difference of 0 for the reference microphone. A time-of-arrival difference for other microphones can then be expressed by a numerical value having either positive or negative polarity since such difference represents either a earlier or later arrival to the microphone in question relative to the reference microphone. If the microphone M1 is chosen as the reference microphone, it follows that time-of-arrival difference signals An(S1fi) - An(S1fn) are all equal to 0.

A sound source status determination unit 111 determines, by a computer operation, any sound source which is not uttering a voice. Initially the time-of-arrival difference signals An(S1F1) -An(S1fn), An(S2f1) -An(S2fn), An(S3f1) - An(S3fn) which are obtained by the band-dependent time difference detector 100 for the common band are compared against each other, thereby determining a channel in which the signal arrives earliest for each band f1 -fn.

For each channel, the total number of bands in which the earliest arrival of the signal has been determined is calculated, and such total number is compared between the channels. As a consequence of this, it can be concluded that the microphone corresponding to the channel having a greater total number of bands is located close to the sound source. If the total number of bands which is calculated for a given channel exceeds a preset reference value ThP, a determination is rendered that there is a sound source in a zone covered by the microphone corresponding to this channel.



BNSDOCID: <EP_____0831458A2_I_>

Levels P(S1) - P(S3) of the respective channels which are detected by the all band level detector 60 are also input to the sound source status determination unit 110. If the level of a particular channel is equal to or less than the preset reference value ThR, a determination is rendered that there is no sound source in a zone covered by the microphone corresponding to that channel.

Assume now that the microphones M1 - M3 are disposed relative to sound sources A, B as illustrated in Fig. 12. It is also assumed that the total number of bands calculated for the channel corresponding to the microphone M1 is denoted by $\chi 1$, and similarly the total numbers of bands calculated for channels corresponding to the microphones M2, M3 are denoted by $\chi 2$, $\chi 3$, respectively.

In this instance, the processing procedure illustrated in Fig. 13 may be used. Specifically, when all of the detection signals P(S1) - P(S3) obtained in the all band level detector 60 are less than the reference value ThR (101), the sound sources A, B are regarded as not uttering a voice, and hence, a control signal SABi is generated (102), thus suppressing both sound source signals SA, SB. At this time, the output signals SA-, SB-represent silent signals.

When only the sound source A is uttering a voice, its sound source signal reaches earliest at the microphone M2 for the frequency components of all the bands, and accordingly the total number of bands $\chi 2$ calculated for the channel corresponding to the microphone M2 is at maximum. When only the sound source B is uttering a voice, its sound source signal reaches the microphone M3 earliest for the frequency components of all the bands, and accordingly, the total number of bands $\chi 3$ calculated for the channel corresponding tot the microphone M3 is at maximum.

When the sound sources A, B are both uttering a voice, the total number of bands in which the sound signal reaches earliest will be comparable between the microphones M2 and M3.

Accordingly, when the total number of bands in which the sound source signal reaches a given microphone earliest exceeds the reference ThP, a determination is rendered that there exists a sound source in a zone which is covered by the microphone, and that that sound source is uttering a voice.

In the above example, when only the sound source A is uttering a voice, only $\chi 2$ exceeds the reference value ThP (see 103 in Fig. 3), providing a detection that the uttering sound source exists in the zone Z3 which is covered by the microphone M2, and accordingly, a control signal SBi is generated (104) to suppress the acoustic signal SB while allowing only the signal SA to be delivered.

When only the sound source B is uttering a voice, only $\chi 3$ exceeds the reference value ThP (105), providing a detection that the uttering sound source exists in the zone Z4 which is covered by the microphone M3, and accordingly, a control signal SAi is generated (106), suppressing the signal SA while allowing only the signal SB to be delivered.

In the present example, ThP is established on the order of n/3, for example, and if the sound sources A, B are both uttering a voice, both $\chi 2$ and $\chi 3$ may exceed the reference value ThP. In such instance, one of the sound sources, which may be the sound source A in the present example, may be given a preference to allow the separated signal corresponding to the sound source A to be delivered, as illustrated by the processing procedure shown in Fig. 13. If both $\chi 2$ and $\chi 3$ are below the reference value ThP, a determination is rendered that both sound sources A, B are uttering a voice as long as the levels P(S1) - P(S3) exceed the reference value ThR, and hence control signals SAi, SBi, SABi are not generated (107 in Fig. 3), thus preventing the suppression of the voice signals SA, SB in the signal suppression unit 90.

When the sound source C is added to the zone Z6 in the arrangement of Fig. 12 as indicated in Fig. 14, the sound source separator 80 delivers a signal SC corresponding to the sound source C, in addition to the signal SA corresponding to the sound source A and the signal SB corresponding to the sound source B, even though this is not illustrated in the drawings. In a corresponding manner, the sound source status determination unit 110 delivers a control signal SCi which suppresses the signal SC in addition to the signal SAi which suppresses the signal SA and a control signal SBi which suppresses the signal SB, and also delivers a control signal SBCi which suppresses the signals SB and SC, a control signal SCAi which suppresses the signal SC and SA, and a control signal SABCi which suppresses all of the signals SA, SB and SC in addition to a control signal SABi which suppresses the signals SA and SB. The operation of the sound source status determination unit 110 remains the same as mentioned previously in connection with Fig. 15.

When all of the levels P(S1) - P(S3) fail to exceed the reference value ThR, a determination is rendered that no sound source A - C is uttering a voice, and the sound source status determination unit 110 delivers a control signal SABCi, thus suppressing all of the signals SA, SB and SC.

When the sound source A, B or C is uttering a voice alone, the time-of-arrival for the channel corresponding to the microphone which is located closest to that sound source will be earliest, in a similar manner as occurs for the two sound sources mentioned above, and accordingly, either one of the total number of bands for the respective channel $\chi 1$, $\chi 2$, $\chi 3$ will exceed the reference value ThP. When only the sound source C is uttering a voice, the control signal SABi is delivered to suppress the signals SA, SB. When only the sound source A is uttering a voice, the control signal SBCi is delivered to suppress the signals SB, SC. Finally, when only the sound source B is uttering a voice, the control signal SACi is delivered to suppress the signals SA, SC (203 - 208 in Fig. 15).

When two of the three sound sources A - C are uttering a voice, the total number of bands which achieved the earliest time-of -arrival for the channel corresponding to the microphone located in a zone in which the non-uttering sound source is disposed will be reduced as compared with the corresponding total numbers for the other microphones. For



BNSDOCID: <EP_____0831458A2_I_>

example, for the sound source C alone is not uttering, the number of bands $\chi 1$ which achieved the earliest time-of-arrival to the microphone M1 will be reduced as compared with the corresponding total numbers of bands $\chi 2$, $\chi 3$ for the remaining two microphones M2, M3.

Accordingly, a preset reference value ThQ (< ThP) is established, and if $\chi 1$ is equal to or less than the reference value ThQ, a determination is rendered with respect to the zones Z5, Z6 divided from the space shared by the microphones M1 and M3 that the sound source located in the zone Z6 which is located close to the microphone M1 is not uttering a voice, and also a determination is rendered with respect to the zones Z1, Z2 divided from the space shared by the microphones M1 and M2 that the sound source in the zone Z1 which is located close to the microphone M1 is not uttering a voice.

In this manner, a determination is rendered that sound sources located within the zones Z1, Z6 are not uttering a voice. Since the sound sources located within these zones represent the sound source C, it follows from these determinations that the sound source C is not uttering a voice. As a consequence, it is determined that only the sound sources A, B are uttering a voice, thus generating the control signal SCi to suppress the signal SC (209 - 210 in Fig. 15). A similar determination is rendered for zones in which either sound source A alone or sound source B alone does not utter a signal (211 - 214 in Fig. 15).

If it is determined that all of $\chi 1$, $\chi 2$, $\chi 3$ are not less than the reference value ThQ, a determination is rendered that all of the sound sources A, B, C are uttering a voice (215 in Fig. 15).

In the above example, the space is divided into six zones Z1 - Z6, but the space can be divided into three zones as illustrated in Fig. 16 where the status of sound sources can also be determined in a similar manner. In this instance, if only the sound source A is uttering a voice, for example, the total number of bands $\chi 2$ for the channel corresponding to the microphone M2 will be at maximum, and accordingly, a determination is rendered that there is a sound source in the zone Z2 covered by the microphone M2. Alternatively, when only the sound source B is uttering a voice, $\chi 3$ will be at maximum, and accordingly, a determination is rendered similarly that there is a sound source in the zone Z3. If $\chi 1$ is equal to or less than the preset value ThQ, a determination is rendered with respect to the zones divided from the space shared by the microphones M1 and M3 that the sound source located within the zone Z1 is not uttering a voice, and similarly a determination is rendered with respect to the zones divided from the space shared by the microphones M1 and M2 that a sound source located within the zone Z1 is not uttering a voice. In this manner, the status of sound sources can be determined when the space is divided into three zones in the same manner as when the space is divided into six zones.

The reference values ThP, ThQ may be established in the same way as when utilizing the band-dependent levels as mentioned above.

While the same reference values ThR, ThP, ThQ are used for all of the microphones M1 - M3, these reference values may be suitably changed for each microphone. While the foregoing description has dealt with the provision of three microphones for three sound sources, the detection of a sound source zone is similarly possible provided the number of microphones is equal to or greater than the number of sound sources. A processing procedure used at this end is similar as when utilizing the band-dependent levels mentioned above. Accordingly, when there are four sound sources, for example, three of which are uttering a voice (or one is silent), the processing may end at this point, but in order to select one of the remaining three sound sources which is close to a silent condition, the reference value may be changed from ThQ to ThS (ThP > ThS > ThQ), and each of the steps 210, 212, 214 shown in Fig. 15 may be followed by a processor section which is constructed in the similar manner as constructed by the steps 209 - 214 shown in Fig. 15, thus determining one of the three sound sources which remains silent.

In the procedure shown in Fig. 17, the time difference may be utilized in place of the level, and in such instance, the processing procedure shown in Fig. 17 is applicable to the suppression of unnecessary signals utilizing the time-of-arrival differences shown in Fig. 18.

The method of separating a sound source according to the invention as applied to a sound collector which is designed to suppress runaround sound will be described. Referring to Fig. 19, disposed within a room 210 is a loudspeaker 211 which reproduces a voice signal from a mate speaker which is conveyed through a transmission line 212, thus radiating it as an acoustic signal into the room 210. On the other hand, a speaker 215 standing within the room 210 utters a voice, the signal from which is received by a microphone 1 and is then transmitted as an electrical signal to the mate speaker through a transmission line 216. In this instance, the voice signal which is radiated from the loudspeaker 211 is captured by the microphone 1 and is then transmitted to the mate speaker, causing a howling.

To accommodate for this, in the present embodiment, another microphone 2 is juxtaposed with the microphone 1 substantially in a parallel relationship with the direction of array of the loudspeaker 211 and the speaker 215, and the microphone 2 is disposed on the side nearer the loudspeaker 211. These microphones 1, 2 are connected to a sound source separator 220. The combination of the microphones 1, 2 and the sound source separator 220 constitutes a sound source separation apparatus as shown in Fig. 1. Specifically, the arrangement shown in Fig. 1 except for the microphones 1, 2 represent a sound separator 220, which is defined more precisely as the arrangement shown in Fig. 1 from which the dotted line frame 9 is eliminated, with the remaining output terminal $t_A$ being connected to the trans-



BNSDOCID: <EP_____0831458A2_I_>

mission line 216. An overall arrangement is shown in Fig. 20, to which reference is made, it being understood that Fig. 20 includes certain improvements.

In the resulting arrangement, the speaker 215 functions as the sound source A shown in Fig. 1 while the loudspeaker 211 serves as the sound source B shown in Fig. 1. As mentioned previously in connection with Fig. 1, the voice signal from the loudspeaker 211 which corresponds to the sound source B is cut off from the output terminal $t_A$ while the voice signal from the speaker 215 which corresponds to the sound source A is delivered alone thereto. In this manner, the likelihood that the voice signal from the loudspeaker 211 is transmitted to the mate speaker is eliminated, thus eliminating the likelihood of a howling occurring.

Fig. 20 shows an improvement of this howling suppression technique. Specifically, a branch unit 231 is connected to the transmission line 212 extending from the mate speaker and connected to the loudspeaker 211, and the branched voice signal from the mate speaker is divided into a plurality of frequency bands in a bandsplitter 233 after it is passed through a delay unit 232 as required. This division may take place into the same number of bands as occurring in the bandsplitter 4 by utilizing a similar technique. Components in the respective bands or band signals from the mate speaker which are divided in this manner are analyzed in transmittable band determination unit 234, which determines whether or not a frequency band for these components lies in a transmittable frequency band. Thus, a band which is free from frequency components of a voice signal from the mate speaker or in which such frequency components are at a sufficiently low level is determined to be a transmittable band.

A transmittable component selector 235 is inserted between the sound source signal selector 602L and the sound source synthesizer 7A. The sound source signal selector 602L determines and selects a voice signal from the speaker 215 from the output signal S1 from the microphone 1, which voice signal is fed to the transmittable component selector 235 where only a component which is determined by the transmittable band determination unit 234 as lying in a transmittable band is selected to the sound source signal synthesizer7A. Accordingly, frequency components which are radiated from the loudspeaker 211 and which may cause a howling can not be delivered to the transmission line 216, thus more reliably suppressing the occurrence of the howling.

The delay unit 232 determines an amount of delay in consideration of the propagation time of the acoustic signal between the loudspeaker 211 and the microphones 1, 2. The delay action achieved by the delay unit 232 may be inserted anywhere between the branch unit 231 and the transmittable component selector 235. If it is inserted after the transmittable band determination unit 234, as indicated by a dotted frame 237, a recorder capable of reading and storing data may be employed to read data at a time interval which corresponds to the required amount of delay to feed it to the transmittable component selector 235. The provision of such delay means may be omitted under certain circumstances.

In the embodiment shown in Fig. 20, components which may cause a howling are interrupted on the transmitting side (output side), but may be interrupted at the receiving side (input side). Part of such embodiment is illustrated in Fig. 21. Specifically, a received signal from the transmission line 212 is divided into a plurality of frequency bands in a bandsplitter 241 which performs a division into the same number of bands as occurring in the bandsplitter 4 (Fig. 1) by using a similar technique. The band splitted received signal is input to a frequency component selector 242, which also receives control signals from the sound source signal determination unit 601 which are used in the sound source signal selector 602L in selecting voice components from the speaker 215 as obtained from the microphone 1. Band components which are not selected by the sound source signal selector 602L, and hence which are not delivered to the transmission line 216, are selected from the band splitted received signal in the frequency component selector 242 to be fed to an acoustic signal synthesizer 243, which synthesizes them into an acoustic signal to feed the loudspeaker 211. The acoustic signal synthesizer 243 functions in the same manner as the sound source signal synthesizer 7A. With this arrangement, frequency components which are delivered to the transmission line 216 are excluded from the acoustic signal which is radiated from the loudspeaker 211, thus suppressing the occurrence of howling.

As mentioned previously in connection with the embodiment shown in Fig. 1, the threshold values $\Delta$Lth, $\Delta\tau$th which are used in determining to which sound source signal the band components belong in accordance with a band-dependent inter-channel time difference or band-dependent inter-channel level difference have preferred values which depend on the relative positions of the sound source and the microphones. Accordingly, it is preferred that a threshold presetter 251 be provided as shown in Fig. 20 so that the thresholds $\Delta$Lth, $\Delta\tau$th or the criterion used in the sound source signal determination unit 601 be changed depending on the situation.

To enhance the noise resistance, a reference value presetter 252 is provided in which a muting standard is established for muting frequency components of levels below a given value. The reference value presetter 252 is connected to the sound source signal selector 602L, which therefore regards the frequency components in the signal collected by the microphone 1 which is selected in accordance with the level difference threshold and the phase difference (time difference) threshold and having levels below a given value as noise components such as a dark noise, a noise caused by an air conditioner or the like, and eliminates these noise components, thus improving the noise resistance.

To prevent the howling from occurring, a howling preventive standard is added to the reference value presetter 252 for suppressing frequency components of levels exceeding a given value below the given value, and this standard is



BNSDOCID: <EP_____0831458A2_I_>

also fed to the sound source signal selector 602L. As a consequence, in the sound source signal selector 602L, those of the frequency components in the signal collected by the microphone 1 which is selected in accordance with the level difference threshold and the phase difference threshold, and additionally in accordance with the muting standard, which have levels exceeding a given value are corrected to stay below a level which is defined by the given value. This correction takes place by clipping the frequency components at the given level when the frequency components momentarily and sporadically exceed the given level, and by a compression of the dynamic range where the given level is relatively frequently exceeded. In this manner, an increase in the acoustic coupling which causes the occurrence of the howling can be suppressed, thus effectively preventing the howling.

An arrangement for suppressing reverberant sound can be added as shown in Fig. 21. Specifically, a runaround signal estimator 261 which estimates a delayed runaround signal and an estimated runaround signal subtractor 262 which is used to subtract the estimated, delayed runaround signal are connected to the output terminal $t_A$. By utilizing the transfer responses of the direct sound and the reverberant sound, the runaround signal estimator 261 estimates and extracts a delayed runaround signal. This estimation may employ a complex cepstrum process which takes into consideration the minimum phase characteristic of the transfer response, for example. If required, the transfer responses of the direct sound and the runaround sound may be determined by the impulse response technique. The delayed runaround signal which is estimated by the estimator 261 is subtracted in the runaround signal subtractor 262 from the separated sound source signal from the output terminal $t_A$ (voice signal from the speaker 215) before it is delivered to the transmission line 216. For a detail of the suppression of the runaround signal by means of the runaround signal estimator 261 and the runaround signal subtractor 262, refer "A.V. Oppenhein and R.W. Schafer 'DIGITAL SIGNAL PROCESSING' PRENTICE-HALL, INC. Press".

Where the speaker 215 moves around only within a given range, a level difference / or a time-of-arrival difference between frequency components in the voice collected by the microphone 1 which is disposed alongside the speaker 215 and frequency components of the voice collected by the microphone 2 which is disposed alongside the loudspeaker 211 are limited in a given range. Accordingly, a criterion range may be defined in the threshold presetter 251 so that signals which lie in the given range of level differences or in a given range of phase difference be processed while leaving the signals lying outside these ranges unprocessed. In this manner, the voice uttered by the speaker 215 can be selected from the signal collected by the microphone 1 with a higher accuracy.

When considered from a different point of view, since the loudspeaker 211 is stationary, a definite level difference and / or phase difference between frequency components of the voice from the loudspeaker 211 which is collected by the microphone 1 disposed alongside the speaker 215 and frequency components for the voice from the loudspeaker 211 which is collected by the microphone 2 disposed alongside it are also limited in a given range. It will be appreciated that such ranges of level difference and phase difference are used as the standard for exclusion in the sound source signal selector 602L. Accordingly, the criterion for the selection to be made in the sound source signal selector 602L may be established in the threshold presetter 251.

When three or more microphones are used in the suppression of the howling, the function of selecting of required frequency components can be defined to a higher accuracy. In addition, while the invention has been described as applied to runaround sound suppressing sound collector of a loudspeaker acoustic system, it should be understood that the invention is also applicable to a telephone transmitter / receiver system as well.

In addition, frequency components which are to be selected in the sound source signal selector 602L are not limited to specific frequency components (voice from the speaker 215) contained in the frequency components of the voice signal which is collected by the microphone 1. Depending on the situation, where an outlet port of an air conditioner system is located toward the speaker 215, for example, it is possible to select those of the frequency components collected by the microphone 2 which are determined as representing the voice of the speaker 215. Alternatively, in an environment having a high noise level, those of the frequency components collected by the microphone 1, 2 which are determined as representing the voice of the speaker 215 may be selected.

The identification of a zone covered by a particular microphone to determine if a sound source located therein is uttering a voice has been described previously with reference to Fig. 12. Thus, it has been described above that it is possible to detect in which one of the zones covered by the microphones M1 - M3 a sound source is located. Thus, when the sound source A is uttering a voice, the total number of bands $\chi 2$ in which the channel corresponding to the microphone M2 exhibits a maximum level is greater than $\chi 1$, $\chi 3$, thus detecting that the sound source A is located within zones Z2, Z3. However, when $\chi 1$ and $\chi 3$ are compared to each other in the arrangement of Fig. 12, it follows that $\chi 1$ is less than $\chi 3$, thus determining that the sound source A is located in the zone Z3. In this manner, the zone of the uttering sound source can be determined to a higher accuracy by utilizing the comparison among $\chi 1$, $\chi 2$, $\chi 3$. Such a comparative detection is applicable to either the use of the band-dependent inter-channel level difference or the band-dependent inter-channel time-of-arrival difference.

In the foregoing description, output channel signals from the microphones are initially subjected to a bandsplitting, but where the band-dependent levels are used, the bandsplitting may take place after obtaining power spectrums of the respective channels. Such an example is shown in Fig.22 where corresponding parts as appearing in Figs. 1 and 11



BNSDOCID: <EP_____0831458A2_I_>

are designated by like reference numerals and characters as before, and only the different portion will be described. In this example, channel signals from the microphones 1, 2 are converted into power spectrums in a power spectrum analyzer 300 by means of the rapid Fourier transform, for example, and are then divided into bands in the bandsplitter 4 in a manner such that essentially and principally a single sound source signal resides in each band, thus obtaining band-dependent levels. In this instance, the band-dependent levels are supplied to the sound source signal selector 602 together with the phase components of the original spectrums so that the sound source signal synthesizer 7 is capable of reproducing the sound source signal.

The band-dependent levels are also fed to the band-dependent inter-channel level difference detector 5 and the sound source status determination unit 70 where they are subject to a processing operation as mentioned above in connection with Figs. 1 and 11. In other respects, the operation remains the same as shown in Figs. 1 and 11.

The method of separating a sound source according to the invention is applied to the suppression of runaround sound or howling has been described above with reference to Figs. 19 to 21. In this howling prevention method / apparatus, the technique of suppressing or muting a synthesized sound from a sound source that is not uttering a voice can also be utilized to achieve a synthesized signal of better quality. A functional block diagram of such an embodiment is shown in Fig. 30 where corresponding parts to those shown in Figs. 1, 11 and Fig. 20 are designated by like reference numerals and characters as used before. Specifically, respective channel signals from microphones 1, 2 are divided each into a plurality of bands in a bandsplitter 4 to feed a sound source signal selector 602L, a band-dependent inter-channel time difference / level difference detector 5 and a band-dependent level / time difference detector 50. Outputs from the microphones 1, 2 are also fed to an inter-channel time difference / level difference detector 3, an inter-channel time difference or level difference from which is fed to the band-dependent inter-channel time difference / level difference detector 5 and to a sound source signal determination unit 601. Output levels from the microphones 1, 2 are fed to a sound source status determination unit 70.

Outputs from the band-dependent inter-channel time difference / level difference detector 5 are fed to the sound source signal determination unit 601 where a determination is rendered as to from which sound source each band component accrues. On the basis of such a determination, a sound source signal selector 602L selects an acoustic signal component from a specific sound source, which is only the voice component from a single speaker in the present example, to feed a sound source signal synthesizer 7. On the other hand, the band-dependent level / time difference detector 50 detects a level or time-of-arrival difference for each band, and such detection outputs are used in the sound source status determination unit 70 in detecting a sound source which is uttering or not uttering a voice. A synthesized signal for a sound source which is not uttering a voice is suppressed in a signal suppression unit 90.

The apparatus operates most effectively when employed to deliver the voice signal from one of a plurality of speakers in a common room who are simultaneously speaking. The technique of suppressing a synthesized signal for a non-uttering sound source can also be applied to the runaround sound suppression apparatus described above in connection with Figs. 20 and 21. The arrangement shown in Fig. 22 is also applicable to the runaround sound suppression apparatus described above in connection with Figs. 19 to 21.

In the embodiment described previously with reference to Fig.2, for each band split signal, it may be determined from which sound source it is oncoming by utilizing only the corresponding band-dependent inter-channel time difference without using the inter-channel time difference. Also in the embodiment described previously with reference to Fig. 5, each band split signal may be determined from which sound source it is oncoming by utilizing the band-dependent inter-channel level difference without using the inter-channel level difference. The detection of the inter-channel level difference in the embodiment described above with reference to Fig. 5 may utilize the levels which prevail before conversion into the logarithmic levels.

It is to be understood that the manner of division into frequency bands need not be uniform among the bandsplitter 4 in Fig. 1, the bandsplitters 40 in Figs. 11 and 18, the bandsplitter 233 in Fig.20 and the bandsplitter 241 in Fig. 21. The number of frequency bands into which each signal is divided may vary among these bandsplitters, depending on the required accuracy. For the sake of subsequent processing, the bandsplitter 233 in Fig. 20 may divide an input signal into a plurality of frequency bands after the power spectrum of the input signal is initially obtained.

It has been described above in connection with the generation of a silent signal suppression control signal with reference to Figs. 11 and 18 that the zone of an uttering sound source can be detected, and that such a detection may be utilized to generate a suppression control signal.

A functional block diagram of an apparatus for detecting a sound source zone according to the invention is shown in Fig. 23 where numerals 40, 50 represent corresponding ones shown by the same numerals in Figs. 11 and 18. Channel signals from the microphones M1 - M3 are each divided into a plurality of bands in bandsplitters 41, 42, 43, and band-dependent level / time difference detectors 51, 52, 53 detect the time-dependent level or time-of-arrival difference for each channel from the band signals in a manner mentioned above in connection with Figs. 11 and 18. These band-dependent level or band-dependent time-of-arrival differences are fed to a sound source zone determination unit 800 which determines in which one of the zones covered by the respective microphones a sound source is located, delivering a result of such a determination.



BNSDOCID: <EP_____0831458A2_I_>

A processing procedure used in the method of detecting a sound source zone will be understood from the flow diagram shown in Fig. 17 and from the above description, but is summarized in Fig. 24, which will be described briefly. Initially, channel signals from the microphones M1 - M3 are received (S1), each channel signal is divided into a plurality of bands (S2), and a level or a time-of-arrival difference of each divided band signal is determined (S3). Subsequently, a channel having a maximum level or of an earliest arrival for the same band is determined (S4). A number of bands which each channel has achieved a maximum level or an earliest arrival, $\chi 1$, $\chi 2$, $\chi 3$, $\cdots$ is determined (S5). A maximum one $\chi_M$ among these numbers $\chi 1$, $\chi 2$, $\chi 3$, $\cdots$ is selected (S6), and a determination is rendered that a sound source is located in a zone covered by a microphone of a channel M which corresponds to $\chi_M$ (S7).

During the selection of $\chi_M$, an examination may be made to see if $\chi_M$ is greater than a reference value, which may be equal to n/3 (where n represents the number of divided bands) (S8) before proceeding to step S7. Subsequent to the step S5, an examination is made (S9) to search for any one of $\chi 1$, $\chi 2$, $\chi 3$, $\cdots$ which exceeds a reference value, which may be 2n/3, for example. If YES, a determination is rendered that there is a sound source in a zone covered by a microphone of the channel M which corresponds to $\chi_M$(S7). To determine the zone with a higher accuracy, when it is found at step S9 that there is a $\chi_M$ which exceeds the reference value, $\chi_{M1}$, $\chi_{M2}$ for channels M1, M2 which are associated with the microphones located adjacent to the microphone for channel M are compared against each other. The sound source zone is determined on the basis of the microphone corresponding to M' for the greater $\chi_{M'}$ (M' being either 1 or 2) and the microphone corresponding to M. Thus, if $\chi_{M1}$ is greater, a determination is rendered that a sound source is located in the zone covered by the microphone for the channel M and located toward the microphone corresponding to M1(S11).

With the method of detecting a sound source zone according to the invention, each microphone output signal is divided into smaller bands, and the level or time-of-arrival difference is compared for each band to determine a zone, thus enabling the detection of a sound source zone in real time while avoiding the need to prepare a histogram.

An experimental example in which the invention comprising a combination of Figs. 6 - 9 is applied will be indicated below. Specifically, the invention is applied to a combination of two sound source signals from three varieties as illustrated in Fig. 25, the frequency resolution which is applied in the bandsplitter 4 is varied, and the separated signals are evaluated physically and subjectively. A mixed signal before the separation is prepared by the addition while applying only an inter-channel time difference and level difference from the computer. The applied inter-channel time difference and level difference are equal to 0.47 ms and 2 dB.

Five values of the frequency resolution including about 5 Hz, 10 Hz, 20Hz, 40 Hz and 80 Hz are used in the bandsplitter 4. An evaluation is made for six kinds of signals including the signals separated according to the respective resolutions and the original signal. It is to be noted that the signal band is about 5 kHz.

A quantitative evaluation takes place as follows: When the separation of mixed signals takes place perfectly, the original signal and the separated signal will be equal to each other, and the correlation coefficient will be equal to 1. Accordingly, a correlation coefficient between original signal and the processed signal is calculated for each sound to be used as a physical quantity representing the degree of separation.

Results are indicated in broken lines 9 in Fig. 27. For any combination of voices, the correlation value is significantly reduced at the frequency resolution of 80 Hz, but no remarkable difference is noted for other resolutions. For bird chirping, no significant difference is noted between the values of frequency resolution used.

A subjective evaluation is made as follows: 5 Japanese men in their twenties and thirties and having a normal audition are employed as subjects. For each sound source, separated sounds at five values of the frequency resolution and the original sound are presented at random diotically through a headphone, asking them to evaluate the tone quality at five levels. A single tone is presented for an interval of about four seconds.

Results are indicated in solid lines in Fig. 27. It is noted that for the separated sound S1, the highest evaluation is obtained for the frequency resolution of 10 Hz. There existed a significant difference ($\alpha < 0.05$) between evaluations for all conditions. As to separated sounds S2 - 4 and 6, the evaluation is highest for the frequency resolution of 20 Hz, but there was no significant difference between 20 Hz and 10 Hz. There existed a significant difference between 20 Hz on one hand and 5 Hz, 40 Hz and 80 Hz on the other hand. From these results, it is found that there exists an optimum frequency resolution independently from the combination of separated voices. In this experiment, a frequency resolution on the order of 20 Hz or 10 Hz represents an optimum value. As to the separated sound S5 (birds chirping), the highest evaluation is given for 40 Hz, but the significant difference is noted only between 40 Hz and 5 Hz and between 20 Hz and 5 Hz. In any instance, there existed a significant difference between the separated sound and the original sound.

Figs. 26 and 28 illustrate the effect brought forth by the present invention.

Fig. 26 shows a spectrum 201 for a mixed voice comprising a male voice and a female voice before the separation, and spectrums 202 and 203 of male voice S1 and female voice S2 after the separation according to the invention. Fig. 28 shows the waveforms of the original voices for male voice S1 and female voice S2 before the separation at A, B, shows the mixed voice waveform at C, and shows the waveforms for male voice S1 and female voice S2 after the separation at D, E, respectively. It is seen from Fig. 26 that unnecessary components are suppressed. In addition, it is seen



BNSDOCID: <EP_____0831458A2_I_>

from Fig. 28 that the voice after the separation is recovered to a quality which is comparable to the original voice.

The resolution for the bandsplitting is preferably in a range of 10 - 20 Hz for voices, and a resolution below 5 Hz or above 50 Hz is undesirable. The splitting technique is not limited to the Fourier transform, but may utilize band filters.

Another experimental example in which the signal suppression takes place in the signal suppression unit 90 by determining the status of the sound source by utilizing the level difference as illustrated in Fig. 11 will be described. A pair of microphones are used to collect sound from a pair of sound sources A, B which are disposed at a distance of 1.5 m from a dummy head and with an angular difference of 90° (namely at an angle of 45° to the right and to the left with respect to the midpoint between the pair of microphones) at the same sound pressure level and in a variable reverberant room having a reverberation time of 0.2 s (500 Hz). Combinations of mixed sounds and separated sounds used are S1 - S4 shown in Fig. 22.

For the separated sounds S1 - S4, the ratio of the number of frames which are determined to be silent to the number of silent frames in the original sound are calculated. As a result, it is found that more than 90% are correctly detected as indicated below.

| | Male (S1) | Female (S2) | Female voice 1 (S3) | Female voice 2 (S4) |
|---|---|---|---|---|
| Detection rate | 99% | 93% | 92% | 95% |

Sounds which are separated according to the fundamental method illustrated in Figs. 5 - 9 and according to the improved method shown in Fig. 11 are presented at random diotically through a headphone, and an evaluation is made for the reduced level of noise mixture and for the reduced level of discontinuity. The separated sounds are S1 - S4 mentioned above, and the subjects are five Japanese in their twenties and thirties and having normal audition. A single sound is presented for an interval of about four seconds, and trials for each sound are three times. As a consequence, the rate at which the reduced level of noise mixture is evaluated is equal to 91.7%for the improved method and is equal to 8.3% for the fundamental method, thus indicating that answers replying that the noise mixture is reduced according to the improved method are considerably higher. However, the evaluation for the detection of discontinuity is equal to 20.3% according to the improved method, and is equal to 80.0% according to the fundamental method, thus indicating that far more replies evaluated that the discontinuities are reduced according to the fundamental method. However, no significant difference is noted between the fundamental and the improved method.

To provide a relative evaluation of the separation performance, a comparison of the degree of separation for five kinds of sounds is made according to the subjective evaluation .

(1) Original sound

(2) Fundamental method (computer): a mixed signal resulting from the addition on the computer while applying an inter-channel time difference (0.47 ms) and a level difference (2 dB) is separated according to the fundamental method;

(3) Improved method (actual environment): a mixed sound collected under the condition used in the experiment to determine a detection rate of silent intervals is separated according to the improved method;

(4) Fundamental method (actual environment): a mixed sound collected under the condition used in the experiment to determine a detection rate of silent intervals is separated according to the fundamental method;

(5) Mixed sound: a axed sound collected under the condition used in the experiment to determine a detection rate of silent intervals.

For the first two axed sounds indicated in the chart of Fig. 25, a total of twenty samples of "mixed sounds" obtained by processing the "original sounds" according to the techniques indicated under the sub-paragraphs (1) - (4) are presented at random diotically through a headphone, and an evaluation of the degree of separation is made at seven levels. A score of 7 is given to "most separated" while a score of 1 is given to the "least separated". The subjects, the interval during which the sounds are presented and the number of trials remain the same as those used during the evaluation of the reduced level of noise mixture.

Results are shown in Fig. 29. Specifically all sound sources (S0) is shown at A, male voice (S1) at B, female voice (S2) at C, female voice 1 (S3) at D, and female voice 2 (S4) at E, respectively. A result of analysis of all the sound sources (S0) and a result of analysis for each variety of sound source (S1) - (S4) exhibited substantially similar tendencies. For all of S0 -S4, the degree of separation increases in the sequence of "(1) original sound", "(2) fundamental method (computer)", "(3) improved method (actual environment)", "(4) fundamental method (actual environment)" and "(5) mixed sound". In other words, the improved method is superior to the fundamental method in the actual method in the actual enviroment.



BNSDOCID: <EP_____0831458A2_I_>

## Claims

1. A method of separating at least one sound source from a plurality of sound sources using a plurality of microphones located as separated from each other, comprising steps of dividing each output channel signal from each microphone into a plurality of frequency bands;

   detecting a difference, between the output channels and for each band, in the value of a parameter in an acoustic signal reaching a microphone which varies attributable to the locations of the plurality of microphones, as band-dependent inter-channel parameter value differences;
   on the basis of the band-dependent inter-channel parameter value differences for respective bands, determining which one of the band divided output channel signals for the respective bands is input from which one of the sound sources, thus determining a sound source signal:
   on the basis of a determination rendered in the sound source signal determining step, selecting at least one of the signals input from a common sound source from the band divided output channel signals;
   and synthesizing a plurality of band signals selected as signals output from the common sound source into a sound source signal.

2. A method according to Claim 1 in which the band division takes place into bands which are chosen small enough to assure that each divided band signal of each output channel signal essentially comprises components of an acoustic signal from a single sound source.

3. A method according to Claim 2 in which the parameter value used in the step of detecting the band-dependent inter-channel parameter value differences comprises a time for an acoustic signal from a sound source to reach each microphone, and in which the band-dependent inter-channel parameter value differences are band-dependent inter-channel time differences which represent differences between the microphones in the time required to reach the respective microphones.

4. A method according to Claim 3, further including the step of detecting differences between the microphones in the time required for the acoustic signal to reach the respective microphones from the output channel signals from the respective microphones, as inter-channel time differences,
   and the step of determining a sound source signal by collating the band-dependent inter-channel time differences to determine from which one of the sound sources the band divided output channel signal of a particular band is input.

5. A method according to Claim 4 in which the step of detecting the inter-channel time differences comprises steps of determining cross-correlations between the output channel signals, and determining the inter-channel time differences as time differences between those output channel signals which exhibit peaks in the cross-correlations.

6. A method according to Claim 5 in which one of the inter-channel time differences which is closest to a time corresponding to a phase difference between components in the same band of the band divided output channels is defined as the band-dependent inter-channel time difference.

7. A method according the Claim 2 in which the parameter value used in detecting the band-dependent inter-channel parameter value differences is a signal level when a acoustic signal from the sound source reaches a microphone, and in which the band-dependent inter-channel parameter value differences represent level differences of the band divided output channels between corresponding bands.

8. A method according to the Claim 7, further comprising the steps of

   detecting level differences between the output channel signal from the respective microphone as inter-channel level differences;
   comparing the inter-channel level differences against all of the corresponding band-dependent inter-channel level differences;
   if a similar relationship applies in the comparing step for a given number or more of the divided bands, determining that the corresponding output channel signal is input from a common sound source for all the bands on the basis of the inter-channel level differences;
   and if the similar relationship is not established for a given number or more of the bands during the comparing step, executing the step of determining the sound source signal in which from which one of the sound sources



BNSDOCID: <EP_____0831458A2_I_>

a signal is input for each band is determined.

9. A method according to Claim 2 in which the parameter value represent a time for an acoustic signal from a sound source to reach a microphone and also represent a signal level when the acoustic signal reaches the microphone, the band-dependent inter-channel parameter value differences being determined as band-dependent inter-channel time differences and as band-dependent inter-channel level differences, further comprising the steps of

detecting differences between the microphones in the time for acoustic signals from the respective sound sources to reach the respective microphones from the output channel signals from the respective microphones, as inter-channel time differences; and dividing the band divided output channel signals into three frequency ranges including a low, a middle and a high range on the basis of the inter-channel time differences; and in which the step of determining a sound source signal comprises the steps of
determining which one of the band-divided output channel signals is input from which one of the sound sources by utilizing the band-dependent inter-channel time differences for the frequency bands in the low range;
determining which one of the band-divided output channel signal is input from which one of the sound sources by utilizing the band-dependent inter-channel level differences and the band-dependent inter-channel time differences for the frequency bands in the middle range;
and determining which one of the band divided output channel signal is input from which one of the sound sources by utilizing the band-dependent inter-channel level differences for frequency bands in the high range.

10. A method according to one of Claims 1 to 9 in which where frequency bands of the original channel signal, between which the band-dependent inter-channel parameter value differences are to be obtained, are different from each other, the step of determining the band-dependent inter-channel parameter value differences is not executed for a frequency band or bands which do not overlap each other, and the band in which the signal is present is determined to be an input signal from a sound source having a previously known broad band in the step of determining a sound source signal.

11. A method of separating at least one sound source from a plurality of sound sources by using a plurality of microphones located as separated from each other, comprising the steps of determining power spectrums for output channel signals from the respective microphones;

dividing the power spectrum of each channel into a plurality of frequency bands so that principally components from a single sound source are contained in each band;
detecting differences in the power spectrums which are divided between the channels and for each common band as band-dependent inter-channel level differences;
on the basis of the band-dependent inter-channel level differences for the respective bands, determining to which one of the output channel signals the signals in a particular band correspond, thus determining a sound source signal;
on the basis of a determination rendered in the step of determining a sound source signal, selecting at least one of the signals from a common sound source on the basis of the divided power spectrum;
and synthesizing the spectrums selected as from the common sound source into a sound source signal.

12. A method according to claim 11, further comprising the steps of

detecting level differences between the output channel signals from the respective microphones as inter-channel level differences;
comparing the inter-channel level differences against all of the corresponding band-dependent inter-channel level differences;
if a similar relationship applies for a given number or more of the divided bands during the comparing step, rendering a determination on the basis of the inter-channel level differences that the output channel signals are input from a common sound source for all the bands,
and if the similar relationship does not apply for the given number or more of the divided bands during the comparing step, executing the step of determining a sound source signal.
dividing in a second bandsplitting process the output channel signals from the respective microphones into a plurality of frequency band chosen such that each bands contains principally components from a single sound source signal.

13. A method according one of the Claims 1 to 10, further comprising the steps of



BNSDOCID: <EP_____0831458A2_I_>

detecting band-dependent levels of the output channel signals which are divided into the bands in the second bandsplitting process;
comparing the band-dependent levels detected during the band-dependent level detecting step between the channels and for the same band, and detecting a sound source which is not uttering a voice based on a result of such comparison, thus determining a status of a sound source;
and suppressing a synthesized signal corresponding to the non-uttering sound source from among the sound source signals which are synthesized during the step of synthesizing the sound source signal in response to a detection signal which indicates the non-uttering sound source.

**14.** A method according to Claim 13 in which the step of determining the status of a sound source comprises the steps of

comparing band-dependent levels between the channels to determine a channel with a highest level for each band,
determining for each channel a total number of bands for which each channel has the highest level,
determining in a first decision step whether of not the number of bands having the highest level exceeds a first reference value,
if it is found at the first decision step that the first reference value is exceeded, estimating the presence of one sound source which is uttering a voice from the location of the microphone for the channel for which the total number of bands having the highest level exceeds the first reference value;
and detecting a sound source or sources other than the estimated sound source as one which is not uttering a voice.

**15.** A method according to Claim 14, further comprising

a second decision step which determines if the total number of bands having the highest level is equal to or less than a second reference value which is less than the second reference value in the event it is determined in the first decision step that the first reference value is not exceeded,
and detecting, if it is determined in the second decision step that the total number of bands is less than the second reference value, a sound source which is not uttering a voice on the basis of the location of the microphone for the channel having a total number of bands of the highest level which is determined to be less than the second reference value.

**16.** A method according to one of Claims 1 to 10, further comprising the steps of

dividing in a second bandsplitting process the output channel signals from the respective microphones into a plurality of frequency band chosen such that each bands contains principally components from a single sound source signal
detecting time-of-arrival differences of the output channel signals to their associated microphones for each band, thus providing band-dependent time differences;
comparing the band-dependent time-of-arrival differences between the channels for each band, and based on a result of such comparison, detecting a sound source which is not uttering a voice;
and suppressing a synthesized signal which corresponds to the non-uttering sound source from among sound source signals which are synthesized in the sound source signal synthesizing step, in response to a detection signal which detected the non-uttering sound source.

**17.** A method according to Claim 3, further comprising the steps of

detecting a sound source which is not uttering a voice on the basis of the result of comparison of the band-dependent inter-channel time differences between the channels for the same band in a step of determining the status of a sound source,
and suppressing a signal corresponding to the non-uttering sound source from among the sound source signals which are synthesized in the step of synthesizing a sound source, in response to a detection signal detecting the presence of non-uttering sound source determined during the step of determining the status of a sound source.

**18.** A method according to Claim 16 or 17 in which the step of determining the status of a sound source comprises the steps of



BNSDOCID: <EP_____0831458A2_I_>

determining a channel in which a sound source signal reached earliest from the comparison of the band-dependent time-of-arrival differences for each band;
determining in a first decision step whether or not a number of bands in which each channel achieved an earliest arrival exceeds a first reference value;
in the event it is determined in the first decision step that the first reference value is exceeded, estimating one sound source which is uttering a voice on the basis of the location of the microphone for the channel which has the number of bands of the earliest arrival exceeding the first reference value;
and detecting a sound source other than the estimated sound source as not uttering a voice.

**19.** A method according to Claim 17 further comprising the steps of

determining in a second decision step, in the event it is determined in the first decision step that the first reference value is not exceeded, if the number of bands of the earliest arrival is below a second reference value which is less than the first reference value;
and in the event it is determined in the second decision step that the number of bands is below the second reference value, detecting one sound source which is not uttering a voice on the basis of the location of the microphone for the channel having the number of bands below the second reference value.

**20.** A method according to Claim 15 or 19 in which the number of sound sources is equal to four or greater, and in which in the event it is determined in the third decision step that the total number of bands of the highest level is less than the third reference value, the third reference value is sequentially incremented consistent with a requirement that the second reference value is not exceeded, thus repeating the same determination as rendered in the third decision step a number of times equal to or less than (M - 2) where M represents the number of sound sources.

**21.** A method according to one of Claims 13 to 20, further comprising the steps of

detecting the level of all frequency components of the output channel signals, thus determining an all band level;
and a third decision step in which an examination is made to see if all of the all frequency component level of the respective channels detected during the all band detecting step are below a third reference value, and transferring to the step of determining the status of a sound source if it is found that some one of the all frequency component levels is not below the third reference value.

**22.** A method according to Claim 21 in which in the event it is determined in the first decision step that the total number of bands of the highest level is equal to or less than the first reference value, all of the synthesized signals for the sound sources which are synthesized in the sound source signal synthesizing step are suppressed.

**23.** A method according to one of Claims 1 to 9, further comprising the steps of

determining a power spectrum for each output channel from the respective microphone,
subjecting the power spectrum of each channel to a division into frequency bands such that components of one sound source are contained principally in one band to detect a band-dependent level,
comparing the band-dependent levels in a common band to determine a channel exhibiting the maximum level for each band,
determining the status of a sound source including determining the number of bands which each channel exhibited the maximum level, determining if the number of such bands exceeds a first reference value, and determining that a sound source or sources other than the sound source in a zone covered by the microphone for the channel for which the number of bands exceeded the first reference value is not uttering a voice,
and suppressing a signal corresponding to the sound source which is determined as not uttering a voice from among the sound source signals which are synthesized in the step of synthesizing a sound source.

**24.** A method according to Claim 23 in which in the event the first reference value is not exceeded, the step of determining the status of a sound source determines whether or not the number of bands in which the highest level is achieved is below a second reference value which is less than the first reference value, and renders a determination that a sound source in a zone covered by the microphone for the channel for which the number of bands is determined to be below the second reference value is not uttering a voice.



BNSDOCID: <EP_____0831458A2_I_>

25. A method according to one of Claims 1 to 24 in which at least one of the sound sources is a speaker while at least one of the other sound sources is electroacoustical transducer means which converts a received signal oncoming from the remote end into an acoustic signal, and in which the step of selecting a sound source signal comprises interrupting components of an acoustic signal from the electroacoustical transducer means which are contained in the band divided channel signal while selecting components of an acoustic signal from the speaker, and transmitting a sound source signal which is synthesized in the step of synthesizing a sound source to the remote end.

26. A method according to Claim 25 further comprising

a second bandsplitting step of dividing a received signal from the electroacoustical transducer means into a plurality of frequency bands according to the same band division scheme as the first mentioned bandsplitting step such that a principal component in each band comprises components of a single sound source signal,
a step of determining a transmittable band determining each band of the band divided received signal as a transmittable band if the level is below a given value,
and a step of selecting a transmittable band in which only those bands in the band signals which are selected in the step of selecting the sound source signal which are determined as being transmittable are selected and fed to the step of synthesizing a sound source.

27. A method according to Claim 26 in which the selection of the transmittable band is delayed in a manner corresponding to a propagation time of an acoustic signal between the electroacoustical transducer means and the microphone.

28. A method according to Claim 25, further comprising,

a second bandsplitting step in which the received signal is divided into a plurality of frequency bands according to the same band division scheme as the first mentioned band division step such that a principal component in each band comprises component of a single sound source signal,
a frequency component selection step in which the band selected in the step of selecting the sound source signal is eliminated from the band divided components of the received signal,
and a re-synthesis step in which the remaining band components of the received signal are synthesized into a signal in the time domain to be fed to the electroacoustical transducer means.

29. A method according to one of Claims 13 to 28 in which the bandsplitting process and the second bandsplitting process occur in a common process.

30. An apparatus for separating at least one sound source from a plurality of sound sources using a plurality of microphones located as separated from each other comprising

bandsplitting means for dividing each output channel signal from the respective microphones into a plurality of frequency bands which are chosen such that essentially and principally components of an acoustic signal from a single sound source are contained in one band;
means for detecting differences, between the band splitted output channel signals for each band, in the value of a parameter in an acoustic signal reaching a microphone which varies as attributable to the locations of the plurality of microphones, as band-dependent inter-channel parameter value differences;
means for determining which one of the band split channels for the respective band is input from which one of the sound sources on the basis of the band-dependent inter-channel parameter value differences, thus determining a sound source signal;
means for selecting at least one of the signals input from a common sound source from the band spilt output channel signals on the basis of a determination rendered in the process of determining a sound source signal;
and means for synthesizing a plurality of band signals which are selected as signals from the common sound source in the process of selecting a sound source signal into a sound source signal.

31. An apparatus according to Claim 30 in which the parameter value used in said means of detecting the band-dependent inter-channel parameter value differences is a time required for an acoustic signal from a sound source to reach each microphone, and the band-dependent inter-channel parameter value differences are differences between the microphones of the time to reach the respective microphones.

32. An apparatus according to Claim 30, further comprising



BNSDOCID: <EP_____0831458A2_I_>

means for detecting differences between the microphones in the time required for the acoustic signal to reach each microphone as inter-channel time differences from the output channel signals from the microphones,
and in which said means for determining a sound source signal comprises means for collating the inter-channel time differences to determine from which one of the sound sources each of the band split output channel signal is input.

**33.** An apparatus according to claim 30 in which the parameter value used in said means for detecting the band-dependent inter-channel parameter value differences is a signal level as an acoustic signal from a sound source reaches each microphone, and the band-dependent inter-channel parameter value differences are band-dependent inter-channel level differences which represent level differences between the band split output channel signals for a corresponding band.

**34.** An apparatus according to claim 33, further comprising means for detecting level differences between the output channel

signals from the microphones as inter-channel level differences,
means for comparing the inter-channel differences against all of the corresponding band-dependent inter-channel level differences,
and means effective, if a similar relationship applies for a given number or more of the split bands in the comparing means, to determine that a corresponding output channel signal is input from a common sound source for all the bands on the basis of the inter-channel level differences, and if a similar relationship does not apply for a given number or more of the split bands in the comparing means, to operate said means for determining a sound source signal for determining, for each band, from which one of the sound sources a signal is input.

**35.** An apparatus according to Claim 30 in which the parameter value represent the time required for an acoustic signal from a sound source to reach the microphone and a signal level as the acoustic signal reaches the microphone, and the band-dependent inter-channel parameter value differences include band-dependent inter-channel time differences and band-dependent inter-channel level differences,

further including means for determining differences between the microphones in the time required from the respective sound sources to reach the respective microphones from output channel signals from the respective microphones, as inter-channel time differences
and range dividing means for dividing the band split output channel signals in three frequency ranges including a low, a middle, and a high range on the basis of the inter-channel time differences,
and in which said means for determining the sound source signal comprises
means effective with the frequency bands in the divided low range to determine which one of the band split output channel signals comprise an input signal from which one of the sound sources by utilizing the band-dependent inter-channel time differences,
means effective with the frequency bands in the divided middle range to determine which one of the band split output channel signals comprises and input signal from which one of the sound sources by utilizing the band-dependent inter-channel level differences and band-dependent inter-channel time differences,
and means effective with the frequency bands in the divided high range to determine which of the band split output channel signals comprises an input signal from which one of the sound sources by utilizing the band-dependent inter-channel level differences.

**36.** An apparatus according to one of claims 30 to 35, further comprising

means for detecting the band-dependent levels of the output channel signals which are subject to the band-splitting process,
means for determining the status of a sound source by comparing the band-dependent levels as detected by the band-dependent level detecting means between the channels for the same band, and detecting a sound source which is not uttering a voice on the basis of a result of such a comparison,
and means for suppressing a signal corresponding to the sound source which is not uttering a voice from among the sound source signals which are synthesized by said means for synthesizing sound source, in response to a detection signal detecting the presence of a sound source which is not uttering a voice as determined by said means for determining the status of the sound source.

**37.** An apparatus according to claim 36, further comprising



BNSDOCID: <EP_____0831458A2_I_>

an all band level detecting means for detecting the levels of all frequency components of the respective output channel signal,
and first decision means for determining if all of the all frequency component levels of the respective channels as detected by the all band level detecting means are below a first reference value, and allowing a transfer to the operation of said means for determining the status of the sound source when any one level is determined to be not below the first reference value.

**38.** An apparatus according to Claim 37 in which said means for determining the status of a sound source comprises

means for comparing the band-dependent level difference between the channels and determining a channel having the highest level for each band,
means for determining a number of bands for which each channel has exhibited the highest level,
second decision means for determining whether or not the number of bands for which the channel exhibited the highest level exceeds a second reference value,
means operative, whenever it is determined by the second decision means that the second reference value is exceeded, to estimate one sound source which is uttering a voice from the location of the microphone for the channel for which the number of bands which the channel achieved the highest level exceeds the second reference value,
and means for detecting a sound source or sources other than the estimated sound source as ones not uttering a voice.

**39.** An apparatus according to claim 37, further comprising,

third decision means operative, in the event it is determined by the second decision means that the second reference value is not exceeded, to determine if the number of bands in which the channel achieved the highest level is below a third reference value which is less than the second reference value,
and means operative, when it is determined by the third decision means that the number of bands is below the third reference value, to detect the presence of one sound source which is not uttering a voice from the location of the microphone for the channel for which the number of bands of the highest level is determined to be below the third reference value.

**40.** An apparatus according to one of the Claims 30 to 35, further comprising

band-dependent time difference detecting means for detecting time-of-arrival differences of the respective band split output channel signals to the microphones for the same band,
means for determining the status of a sound source for detecting the presence of a sound source which is not uttering a voice on the basis of a result of comparison of the band-dependent time-of-arrival differences as detected by the band-dependent time difference detecting means between the channels and for the same band,
and means for suppressing a signal corresponding to a sound source which is not uttering a voice from among the sound source signals which are synthesized by the sound source synthesizing means, in response to a detection signal detecting the presence of a sound source not uttering a voice which is determined by said means for determining the status of a sound source.

**41.** An apparatus according to Claim 40, further comprising all band level detecting means for detecting the levels of all

frequency components of the respective output channel signals,
and first decision means for determining if all of the all frequency component levels of the respective channels as detected by the all band level detecting means are below a first reference value, and allowing a transfer to the operation of said means for determining the status of a sound source when any one level is determined to be not below the first reference value.

**42.** An apparatus according to Claim 41 in which said means for determining the status of a sound source comprises

means for determining for each band a channel in which the earliest arrival of a sound source signal is achieved from the comparison of the band-dependent time-of-arrival differences,
second decision means for determining if a number of bands in which each channel has achieved the earliest arrival exceeds a second reference value,



BNSDOCID: <EP_____0831458A2_I_>

means operative, whenever it is determined by the second decision means that the second reference value is exceeded, to estimate one sound source which is uttering a voice from the location of the microphone for the channel for which the number of bands achieving the earliest arrival exceeds the second reference value,

and means for detecting a sound source or sources other than the estimated sound source as ones not uttering a voice.

**43.** An apparatus according to Claim 42, further comprising third decision means operative, whenever it is determined by the second decision means that the second reference value is not exceeded, to determine if the number of bands in which the earliest arrival is achieved is below a third reference value which is less than the second reference value,

and means operative, whenever it is determined by the third decision means that the number of bands is below the third reference value, to detect one sound source which is not uttering a voice from the location of the microphone for the channel for which the number of bands is determined to be below the third reference value.

**44.** An apparatus according to one of the Claims 30 to 43 in which at least one of the sound sources is a speaker while at least one of the other sound sources is an electroacoustical transducer means which converts a received signal oncoming from the remote end into an acoustic signal, and in which said means for selecting the sound source signal comprises means for interrupting components of acoustic signal from the electroacoustical transducer means contained in the band split channel signals while selecting components of an acoustic signal from the speaker, further comprising

means for transmitting a sound source signal which is synthesized by the sound source synthesizing means to the remote end.

**45.** An apparatus according to Claim 44, further comprising

a second bandsplitting means for dividing the received signal from the electroacoustical transducer means into a plurality of frequency bands according to the same band division scheme as the first mention bandsplitting means such that only components from a single sound source signal are principally contained in one band,

means for determining a transmittable band for each band of the band divided received signal when its level is below a given value,

and means for selecting only those bands in the band signals which are selected by the sound source signal selecting means as being transmittable and feeding them to the sound source synthesizing process.

**46.** An apparatus according to Claim 45 in which the selection by the transmittable band selecting means is delayed in a manner corresponding to a propagation time of an acoustic signal between the electroacoustical transducer means and the microphone.

**47.** An apparatus according to Claim 44, further comprising

second bandsplitting means for dividing the received signal into a plurality of frequency bands according to the same band division scheme as in the first mentioned bandsplitting means such that principally components from a common sound source are contained in one band;

frequency component selecting means for eliminating the bands which are selected by the sound source signal selecting means from the band divided components of the received signal,

and re-synthesis means for synthesizing remaining band components in the received signal into a signal in the time domain and feeding it to the electroacoustical transducer means.

**48.** An apparatus according to one of Claims 30 to 47 , further comprising threshold presetting means which selects a criterion to be used in said means for determining the sound source signal.

**49.** An apparatus according to one of Claims 30 to 48, further comprising means for establishing a reference value which is used for excluding the band-dependent inter-channel parameter value differences which are above the reference value from the determination.

**50.** An apparatus according to one of Claims 30 to 49 in which said means for selecting the sound source signal comprises reference value presetting means which presets a criterion for muting band components of levels below a given value.



BNSDOCID: <EP_____0831458A2_I_>

51. An apparatus according to one of Claims 30 to 50, further comprising subtracting means for subtracting a delayed run around signal from the synthesized signal supplied from the sound source signal synthesizing means.

52. A record medium having recorded therein a program for a method for separating at least one sound source from a plurality of sound sources using a plurality of microphones located as separated from each other, the method comprising the steps of

dividing each of output channel signals from the microphones into a plurality of frequency bands chosen such that essentially and principally components of an acoustic signal from a single sound source is contained in one band;
detecting differences, between the band divided output channel signals and for each band, in the value of a parameter in an acoustic signal reaching a microphone which varies as attributable to the locations of the plurality of microphones, as band-dependent inter-channel parameter value differences;
on the basis of the band-dependent inter-channel parameter value differences of the respective bands, determining which one of the band divided output channel signals for the respective band is input from which one of the sound sources, thus determining a sound source signal;
selecting at least one of the signals input from a common sound source from the band divided output channel signals on the basis of a determination rendered in the process of determining a sound source signal;
and synthesizing a plurality of band signals which are selected as signals from the common sound source in the process of selecting a sound source signal into a sound source signal.

53. A record medium according to Claim 52 in which the parameter value used in the process of detecting the band-dependent inter-channel parameter value differences is the time required for an acoustic signal from a sound source to reach each microphone, the band-dependent inter-channel parameter value differences are band-dependent inter-channel time differences which represent differences between the microphones in the time required to reach each microphone.

54. A record medium according to claim 53 in which the method comprises a step of

detecting differences between the microphones in the time for an acoustic signal to reach each microphone from the output channel signals of the microphones as inter-channel time differences, and in which the step of determining a sound source signal collates the inter-channel time differences from the band-dependent inter-channel time differences and determines from which one of the sound sources each of the band divided output channel signals of the respective bands is input.

55. A record medium according to Claim 54 in which the step of detecting the inter-channel time differences includes obtaining the cross-correlations between the respective output channel signals, and determining the inter-channel time differences as differences between the output channel signals where the cross-correlations exhibit respective peaks.

56. A record medium according to Claim 55 in which the band-dependent inter-channel time differences are determined by obtaining one close to a time which corresponds to phase differences between components of the band divided output channels for the single band.

57. A record medium according to claim 52 in which the parameter value used in the step of detecting the band-dependent inter-channel parameter value differences are signal levels as acoustic signals from the sound sources reach the respective microphones, and the band-dependent inter-channel parameter value differences are band-dependent inter-channel level differences which represent level differences between corresponding bands of the band divided output channel signals.

58. A record medium according to Claim 57 in which the method further comprises

a step of detecting level differences between the output channel signals from the microphones as inter-channel level differences,
a step of comparing the inter-channel level differences against all of the band-dependent inter-channel level differences,
and when a similar relationship applies for a given number or more of the divided bands in the comparing step,
a step of determining a corresponding output channel signal as being input from common sound source for all



BNSDOCID: <EP_____0831458A2_I_>

the bands on the basis of the inter-channel level differences, and if a similar relationship does not apply for a given number or more of the divided bands in the comparing step, a step of determining from which one of the sound sources the signal is input for respective band, thus executing the step of determining a sound source signal.

**59.** A record medium according to Claim 52 in which the parameter value represent a time required for an acoustic signal from a sound source to reach the microphone and a signal level as the acoustic signal reaches the microphone, and the band-dependent inter-channel parameter value differences include band-dependent inter-channel time differences and band-dependent inter-channel level differences, the method further comprising

a step of detecting differences between the microphones in the time required for the acoustic signal from each sound source to reach the respective microphones from the output channel signals from the microphones as inter-channel time differences,
and a step of dividing the band divided output channel signals into three frequency ranges including a low, a middle and a high range on the basis of the inter-channel time differences,
and in which the step of determining a sound source signal comprises the steps of
determining, for the frequency bands in the divided low range, which one of the band divided output channel signals comprises an input signal from which one of the sound sources by utilizing the band-dependent inter-channel time differences,
determining, for the frequency bands in the divided middle range, which one of the band divided output channel signals comprises an input signal from which one of the sound sources by utilizing the band-dependent inter-channel level differences and the band-dependent inter-channel time differences,
and determining, for the frequency bands in the divided high range, which one of the band divided output channel signals comprises an input signal from which one of the sound sources by utilizing the band-dependent inter-channel level differences.

**60.** A record medium according to one of Claims 52 to 59 in which the method comprises further steps of

detecting a band-dependent level of each of the band divided output channel signals,
determining the status of a sound source by comparing the band-dependent levels between the channels for the same band and detecting a sound source which is not uttering a voice on the basis of the result of such a comparison,
and suppressing a signal which corresponds to the sound source which is not uttering a voice from among the sound source signals which are synthesized in the step of synthesizing the sound source, in response to a detection signal detecting the presence of a sound source which is not uttering a voice and which is obtained in the step of determining the status of a sound source.

**61.** A record medium according to Claim 60 in which the method further comprises

a step of detecting levels of all frequency components of the respective output channel signals to provide an all band level,
and a first decision step of determining if all of the all frequency component levels of the respective channels as detected in the step of detecting the all band level are below a first reference value, and allowing a transfer to the step of determining the status of a sound source whenever anyone of the levels is determined not to be below the first reference value.

**62.** A record medium according to Claim 61 in which the step of determining the status of a sound source comprises the steps of

comparing the band-dependent levels between the channels to determine a channel having the highest level for each band,
determining a number of bands for which each channel has exhibited the highest level,
determining in a second decision step whether or not the number of bands determined exceeds the second reference value,
if it is determined in the second decision step that the second reference value is exceeded, estimating one sound source which is uttering a voice from the location of the microphone for the channel for which the number of bands exceeded the second reference value,
and detecting a sound source or sources other than the estimated sound source as ones not uttering a voice.



BNSDOCID: <EP_____0831458A2_I_>

**63.** A record medium according to Claim 61 in which the method further comprises

a third decision step of determining, whenever it is determined in the second decision step that the second reference value is not exceeded, if the number of bands which exhibit the highest level is below a third reference value which is less than the second reference value,
and if it is determined at the third decision step that the number of bands is below the third reference value, a step of detecting one sound source which is not uttering a voice from the location of the microphone for the channel for which the number of bands is determined to be below the third reference value.

**64.** A record medium according to claim 63 in which there are three or more sound sources, and when it is determined in the third decision step that the number of bands is below the third reference value, the third reference value is sequentially incremented consistent with the requirement that the second reference value is not exceeded to repeat the same process as the third decision step (M - 2 ) times where M represents the number of sound sources.

**65.** A record medium according to one of Claims 52 to 59 in which the method further comprises

a step of detecting band-dependent time differences in which time-of-arrival differences of the respective band divided output channel signals to the microphones are detected for each band,
a step of determining the status of a sound source in which the band-dependent time-of-arrival differences are compared between the channels for the same band, and a sound source not uttering a voice is detected on the basis of the result of such a comparison,
and a step of suppressing a signal corresponding to the sound source which is not uttering a voice from among the sound source signals which are synthesized in the step of synthesizing a sound source in response to a detection signal detecting the presence of a sound source which is not uttering a voice and which is determined in the step of determining the status of a sound source.

**66.** A record medium according to claim 65 in which the method further comprises

a step of detecting all band level in which levels of all frequency components of the respective output channel signals are detected,
and a first decision step of determining if all of the all frequency component levels of the channels are below a first reference value, and if any one level is determined to be not below the first reference value, allowing a transfer to the step of determining the status of a sound source.

**67.** A record medium according to the Claim 66 in which the step of determining the status of a sound source comprises

a step of determining, for each band, a channel in which the sound source signal reached earliest from the comparison of the band-dependent time-of-arrival differences,
a second decision step of determining if a number of bands which each channel achieved an earliest arrival exceed a second reference value,
if it is determined in the second decision step that the second reference value is exceeded, a step of estimating one sound source which is uttering a voice from the location of the microphone for the channel for which the number of bands exceeded the second reference value,
and a step of detecting a sound source of sources other than the estimated one as ones not uttering a voice.

**68.** A record medium according to Claim 67 in which the method further comprises

if it is determined in the second decision step that the second reference value is not exceeded, a third decision step of determining if the number of bands for the earliest arrival is below a third reference value which is less than the second reference value,
and if it is determined at the third decision step that the number of bands is below the third reference value, a step of detecting one sound source which is not uttering a voice from the location of the microphone for the channel for which the number of bands is determined as being below the third reference value.

**69.** A record medium according to Claim 68 in which there are four or more sound sources, and when it is determined in the third decision step that the number of bands is below the third reference value, the third reference value is sequentially incremented consistent with a requirement that the second reference value is not exceeded to repeat



BNSDOCID: <EP_____0831458A2_1_>

the same determination as made in the third decision step a number of times equal or less than (M-2) where M represents the number of sound sources.

**70.** A record medium according to Claim 53 in which the method further comprises

a step of determining the status of a sound source in which band-dependent inter-channel time differences are compared between the channels for the same band and a sound source not uttering a voice is detected on the basis of a result of such a comparison,
and a step of suppressing a signal corresponding to the sound source which is not uttering a voice from among the sound sources signals which are synthesized in the step of synthesizing a sound source signal, in response to a detection signal detecting the presence of a sound source not uttering a voice and obtained in the step of determining the status of a sound source.

**71.** A record medium according to Claim 70 in which the method further comprises

an all band level detecting step in which levels of all frequency components of the respective output channel signals are detected,
and a first decision step to determine if all of the all frequency component levels of the respective channels as detected in the all band level detecting step are below a first reference value, and allowing a transfer to the step of determining the status of a sound source if any one of them is determined to be not less than the first reference value.

**72.** A record medium according to Claim 71 in which the step of determining the status of a sound source comprises the steps of

determining, for each band, a channel in which the sound source signal reached earliest from the comparison of the band-dependent inter-channel time differences,
a second decision step for determining whether or not the number of bands which each channels achieved the earliest arrival exceed a second reference value,
if it is determined in the second decision step that the second reference value is exceeded, estimating one sound source which is uttering a voice from the location of the microphone for the channel for which the number of bands exceeded the second reference value,
and detecting a sound source or sources other than the estimated sound source as ones not uttering a voice.

**73.** A record medium according to Claim 72 in which the method further comprises

if it is determined at the second decision step that the second reference value is not exceeded, a third decision step of determining whether or not the number of bands for the earliest arrival is below a third reference value which is less than the second reference value, and if it is determined at the third decision step that the number of bands is below the third reference value, detecting one sound source which is not uttering a voice from the location of the microphone for the channel for which the number of bands is determined to be below the third reference value.

**74.** A record medium according to one of Claims 52 to 59 in which at least one of the sound sources is a speaker while at least one of the other sound sources is electroacoustical transducer means which transduces a received signal oncoming from the remote end into an acoustic signal, and in which said components of an acoustic signal from the electroacoustical transducer means contained in the band divided channel signals are interrupted while components of an acoustic signal from the speaker are selected, further comprising the step of

transmitting a sound source signal which is synthesized in the step of synthesizing a sound source signal to the remote end.

**75.** A record medium according to Claim 74 in which the method further comprises

a second bandsplitting step for dividing the received signal from the electroacoustical transducer means into a plurality of frequency bands according to the same band division scheme as the first mentioned band division step,
a step of determining a transmittable band for each band of the band divided received signal when its level is



BNSDOCID: <EP_____0831458A2_I_>

below a given value,
and a step of selecting a transmittable band in which only those bands in the band signals as selected in the step of selecting the sound source signal which are determined as being transmittable are selected and fed to the step of synthesizing the sound source.

**76.** A record medium according to Claim 75 in which the selection of the transmittable bands are delayed in a manner corresponding to the propagation time of an acoustic signal between the electroacoustical transducer means and the microphone.

**77.** A record medium according to Claim 72 in which the method further comprises

a second bandsplitting step in which the received signal is divided into a plurality of frequency bands according to the same band division scheme as the first mentioned bandsplitting step,
a step of selecting frequency components in which the bands selected in the step of selecting the sound source signal are eliminated from the band divided components of the received signal,
and a re-synthesis step in which the remaining band components of the received signal are synthesized into a signal in the time domain to be fed to the electroacoustical transducer means.

**78.** A method of detecting a sound source zone in which a zone in which a sound source is located is determined by using a plurality of microphones which are located as separated as from each other, comprising the steps of

dividing each of the output channel signals from the microphones into a plurality of frequency bands, and detecting a parameter value in an acoustic signal reaching a microphone for each band of the band divided output channel signals as band-dependent parameter values, the parameter values undergoing a change as attributable to the location of the plurality of microphones,
and comparing the parameter values detected between the channels for each band and determining a zone in which a sound source for the acoustic signal which is input to the microphone is located on the basis of the result of such comparison.

**79.** A method according to Claim 78 in which the division into bands comprises a small subdivision chosen such that a divided band signal for each output channel signal principally comprises components of an acoustic signal from a single sound source.

**80.** A method according to Claim 79 in which the parameter represents a level of the acoustic signal, and in which the step of determining a sound source zone comprises determining a channel which exhibited a highest level during a comparison of the levels between channels, determining the number of bands for which each channel exhibited the highest level, and determining a zone covered by the microphone for the channel which exhibited the maximum number of bands having the highest level as a sound source zone.

**81.** A method according to Claim 80 in which the step of determining the sound source zone determines a sound source zone covered by the microphone for the channel for which the number of bands having the highest level is at maximum and for which the number of bands is equal to or greater than a reference value.

**82.** A method according to claim 79 in which the parameter represents a level of the acoustic signal, and the step of determining a sound source zone comprises determining a channel which exhibits a highest level by a comparison of the levels between the channels, determining a number of bands for which each channel exhibited the highest level, and determining a zone covered by the microphone for the channel for which the number of bands having the highest level is equal to or greater than a reference value as a sound source zone.

**83.** A method according to Claim 82 in which the number of microphones is equal to three or more, and further comprising the steps of comparing a number of bands having the highest level for each channel other than the channel for which the number of bands exceeds the reference value, and more accurately determining the sound source zone from the zone covered by the microphone for the channel having a greater number of bands having the highest level and a zone covered by the microphone for which the number of bands exceeds the reference value.

**84.** A method according to Claim 78 in which the parameter represents a time-of-arrival differences between the channels, and in which the step of determining the sound source zone comprises determining a channel of the earliest arrival as determined from the comparison of a time-of-arrival differences between the channels, determining a



BNSDOCID: <EP_____0831458A2_I_>

number of bands for which each channel achieved the earliest arrival, and determining a zone covered by the microphone for the channel for which the number of bands having achieved the earliest arrival is at maximum as a sound source zone.

85. A method according to Claim 84 in which the step of determining a sound source zone comprises determining a sound source covered by the microphone for the channel for which the number of bands having achieved the earliest arrival is at maximum and for which the number of bands is equal to or greater than a reference value.

86. A method according to Claim 78 in which the parameter represents a time-of-arrival differences between the channels, and in which the step of determining a sound source zone comprises determining a channel in which the earliest arrival is achieved as determined from the comparison of the time-of-arrival differences between the channels, determining a number of bands for which each channel achieved the earliest arrival, and determining a zone covered by the microphone for the channel for which the number of bands having the earliest arrival is equal to or greater than a reference value as a sound source zone.

87. A method according to claim 86 in which the number of microphones is equal to three or more, further comprising the steps of comparing a number of bands having achieved the earliest arrival for each of the channels other than the channel for which the number bands is equal to or greater than the reference value, and more accurately determining the sound source zone from a zone covered by the microphone for the channel having a greater number of bands having achieved the earliest arrival and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.

88. A method of detecting a sound source zone in which a zone in which a sound source is located is determined by using a plurality of microphones located as separated from each other, comprising

spectrum transform step of transforming an output channel signal from each microphone into a power spectrum,
a bandsplitting step for dividing the power spectrum for each channel into a plurality of bands in a manner such that each band principally contains only signal components from a sound source, thus deriving a level for each band,
a step of comparing the levels between the channels for each divided band to determine a channel which has a maximum level in each band,
a step of determining a number of bands having the maximum level for each channel,
and a step of determining a zone covered by the microphone for the channel for which the number of bands having the highest level is equal to or greater than a reference value as a sound source zone.

89. A method according to Claim 88 in which the number of microphones is equal to three or more, further comprising the steps of comparing a number of bands having the maximum level for each channel other than the channel for which the number of bands is equal to or greater than a reference value, and more accurately determining the sound source zone from a zone covered by the microphone for the channel having a greater number of bands having the highest level and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.

90. An apparatus for detecting a sound source zone in which a zone in which a sound source is located is determined by using a plurality of microphones located as separated from each other, comprising

bandsplitting means for dividing each of output channel signals from respective microphones into a plurality of frequency bands chosen such that one band principally contains only components of an acoustic signal from a single sound source,
means for detecting the value of a parameter in an acoustic signal reaching a microphone for each common band of the respective output channel signals which are divided by the bandsplitting means as band-dependent parameter values which undergo a change as attributable to the location of the plurality of microphones,
and means for comparing the parameter values between the channels for each band and determining a zone in which a sound source for the acoustic signal which is input to the microphone is located on the basis of a result of such comparison.

91. An apparatus according to Claim 90 in which the parameter represents a level of the acoustic signal, and the means for determining a sound source zone comprises means for determining a channel having a highest level as



BNSDOCID: <EP_____0831458A2_I_>

determined from the comparison of levels between the channels, means for determining the number of bands for which each channel exhibited the highest level, and means for determining the zone covered by the microphone for the channel for which the number of bands is at maximum as a sound source zone.

92. An apparatus according to Claim 90 in which the parameter represents a level of the acoustic signal, and the means for determining a sound source zone comprises means for determining a channel which exhibits a highest level as determined from a comparison of the levels between the channels, means for determining a number of bands for which each channel exhibits the highest level, and means for determining a zone covered by the microphone for the channel for which the number of bands is equal to or greater than a reference value as a sound source zone.

93. An apparatus according to Claim 92 in which the number of microphones is equal to three or more, and further comprising comparison means for comparing a number of bands for which each channel other than the channel for which the number of bands is equal to or greater than a reference value exhibits a highest level, and means for more accurately determining a sound source zone from a zone covered by the microphone for the channel having a greater number of bands of the highest level and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.

94. An apparatus according to Claim 90 in which the parameter represents a time-of-arrival difference of the acoustic signal, and in which the means for determining a sound source zone comprises means for determining a channel in which the earliest arrival is achieved as determined from the comparison of the time-of-arrival differences between the channels, means for determining a number of bands for which each channel achieved the earliest arrival, and means for determining a zone covered by the microphone for the channel for which the number of bands achieving the earliest arrival is at maximum as a sound source zone.

95. An apparatus according to Claim 90 in which the parameter represents a time-of-arrival difference of the acoustic signal, further comprising band-dependent time difference detecting means in which time-of-arrival differences between the channels are detected for each band, and in which the means for determining a sound source zone comprises means for determining a channel in which the earliest arrival is achieved as determined from the comparison of the time-of-arrival differences between the channels, means for determining a number of bands for which each channel has achieved the earliest arrival, and means for determining a zone covered by the microphone for the channel for which the number of bands having achieved the earliest arrival is equal to or greater the a reference value as a sound source zone.

96. Apparatus according to claim 90 in which the number of microphones is equal to three or more, further comprising comparison means for comparing a number of bands achieving the earliest arrival between the channels other than the channel for which the number of bands is equal to or greater than a reference value, and means for more accurately determining a sound source zone from a zone covered by the microphone for the channel having a greater number of bands having achieved the earliest arrival and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.

97. A record medium having recorded therein a program for a method of detecting a sound source zone in which a zone in which a sound source is located is determined by using a plurality of microphones located as separated from each other, the method comprising

a step of dividing each of output channel signals form the microphones into frequency bands chosen such that one band principally contains only components of an acoustic signal from a single sound source, and detecting the value of a parameter in an acoustic signal reaching a microphone for each common band of respective output signals which are divided in the band dividing step, thus providing band-dependent parameter values which undergo a change as attributable to the location of the plurality of microphones,
and a step of determining a sound source zone in which the parameter values detected are compared between the channels for each band, and determining a zone in which a sound source for the acoustic signal which is input to the microphone is located is determined on the basis of result of such comparison.

98. A record medium according to Claim 97 in which the parameter represents the level of acoustic signal, and the step of determining a sound source zone comprises determining a channel which exhibits a highest level in the comparison of levels between the channels, determining a number of bands for which each channel has exhibited the highest level, and determining a zone covered by the microphone for the channel for which the number of bands is at



BNSDOCID: <EP_____0831458A2_I_>

maximum as a sound source zone.

99. A record medium according to Claim 98 in which the step of determining the sound source zone determines the sound source zone as a zone covered by the microphone for the channel for which the number of bands having the highest level is at maximum and the number of bands is equal to or greater than a reference value.

100. A record medium according to Claim 97 in which the parameter represents a level of the acoustic signal, and the step of determining a sound source zone comprises determining a channel which exhibits a highest level as determined from the comparison of the levels between the channels, determining a number of bands for which each channel has exhibited the highest level, and determining a zone covered by the microphone for the channel for which the number of bands having the highest level is equal to or greater than a reference value as a sound source zone.

101. A record medium according to Claim 100 in which the number of microphones is equal to three or more, further comprising the steps of comparing a number of bands exhibiting the highest level between channels other than the channel for which the number of bands is equal to or greater than a reference value, and more accurately determining a sound source zone from a zone covered by the microphone for the channel having a greater number of bands exhibiting the highest level and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.

102. A record medium according to Claim 97 in which the parameter represents a time-of-arrival difference of the acoustic signal, the step of determining a sound source zone comprising determining a channel which achieved the earliest arrival as determined from a comparison of the time-of-arrival differences between the channels, determining a number of bands achieving the earliest arrival for each channel, and determining a zone covered by the microphone for the channel for which the number of bands achieving the earliest arrival is at maximum as a sound source zone.

103. A record medium according to Claim 102 in which the step of determining a sound source zone determines a sound source zone as a zone covered by the microphone for the channel for which the number of bands achieving the earliest arrival is at maximum and the number of bands is equal to or greater than a reference value.

104. A record medium according to Claim 97 in which the parameter represents a time-of-arrival difference of the acoustic signal, the step of determining a sound source zone comprising determining a channel achieved the earliest arrival as determined by the comparison of the time-of-arrival differences between the channels, determining a number of bands in which the earliest arrival is achieved for each channel, and determining a zone covered by the microphone for the channel for which the number of bands achieving the earliest arrival is equal to or greater than a reference value.

105. A record medium according to Claim 104 in which the number of microphones is equal to three or more, further comprising the steps of comparing a number of bands achieved the earliest arrival by respective channels other than the channel for which the channel for which the number of bands is equal to or greater than a reference value, and more accurately determining the sound source zone from a zone covered by the microphone for the channel having a greater number of bands achieving the earliest arrival and a zone covered by the microphone for the channel for which the number of bands exceeds the reference value.



BNSDOCID: <EP_____083145BA2_I_>

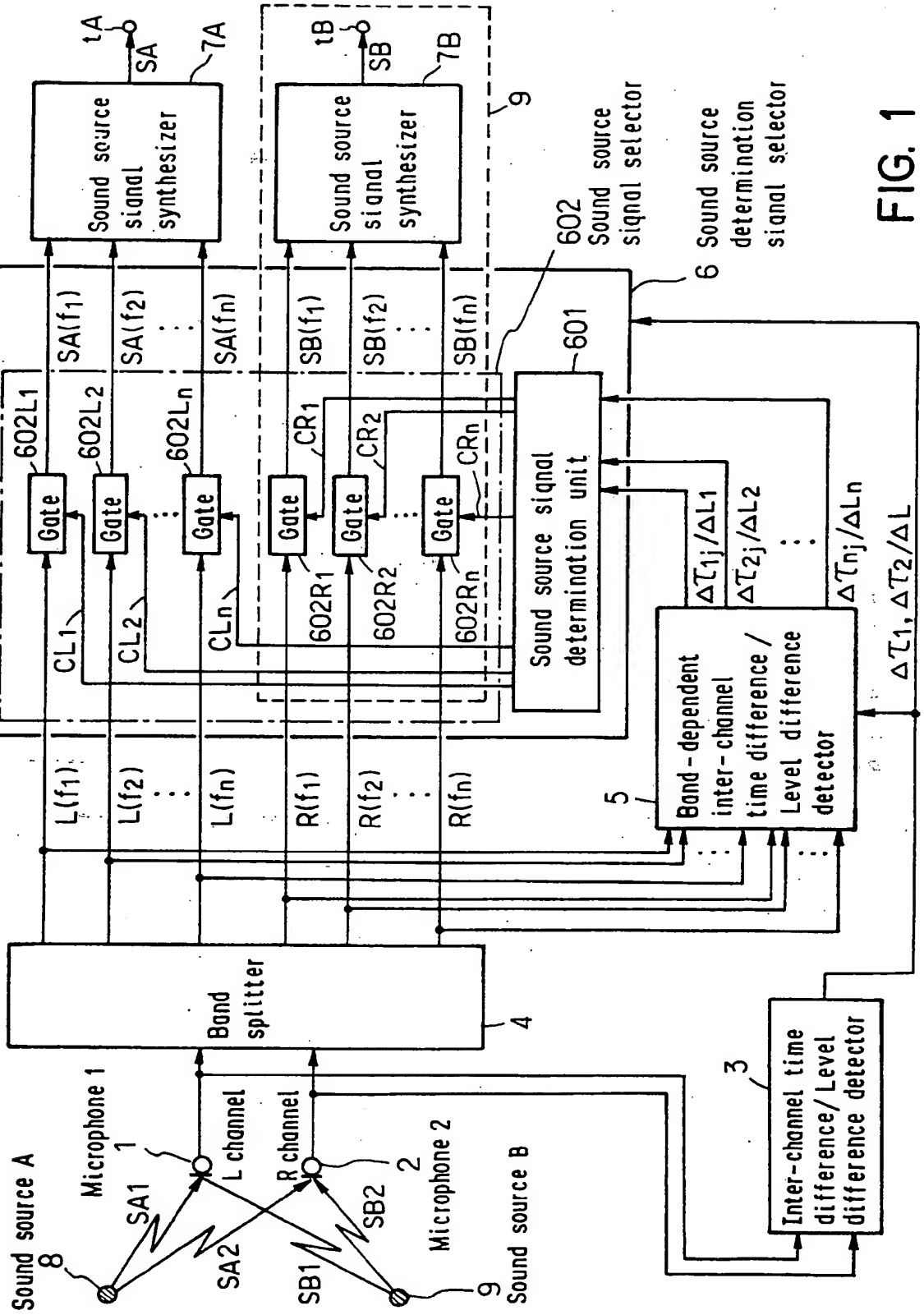



FIG. 1



BNSDOCID: <EP_____0831458A2_I_>

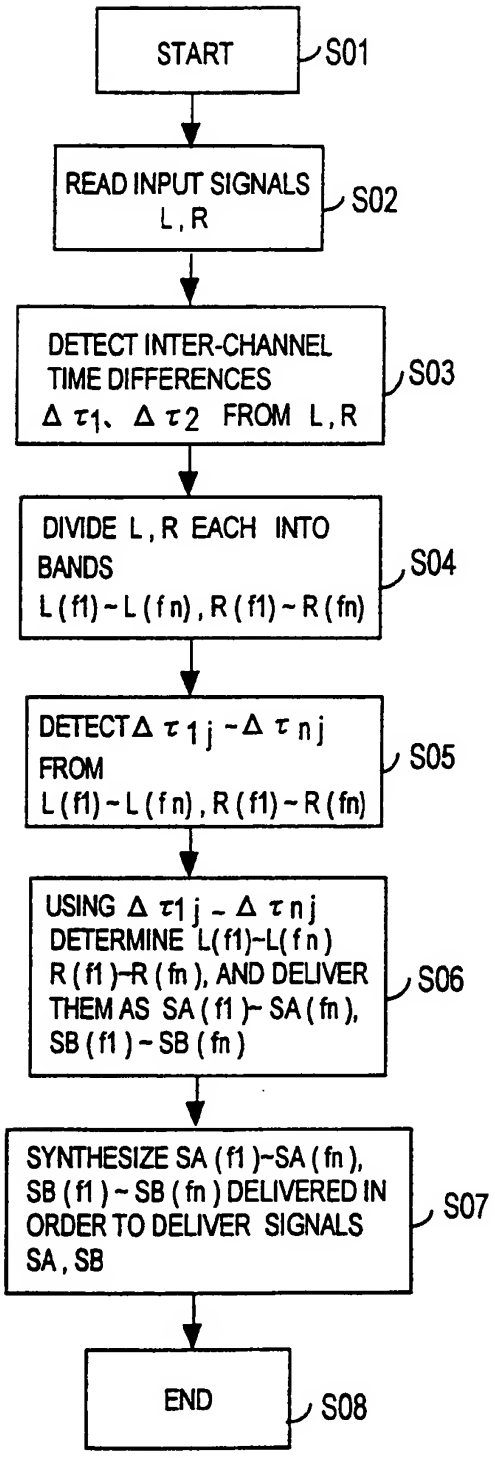



FIG. 2



BNSDOCID: <EP_____0831458A2_I_>

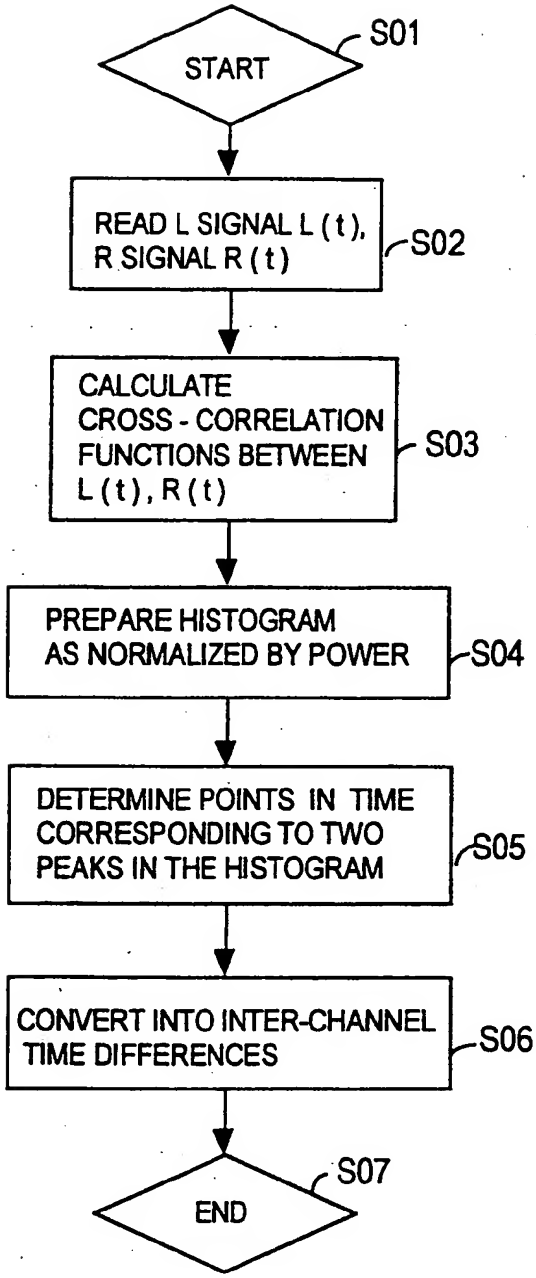



FIG.3



BNSDOCID: <EP_____0831458A2_I_>

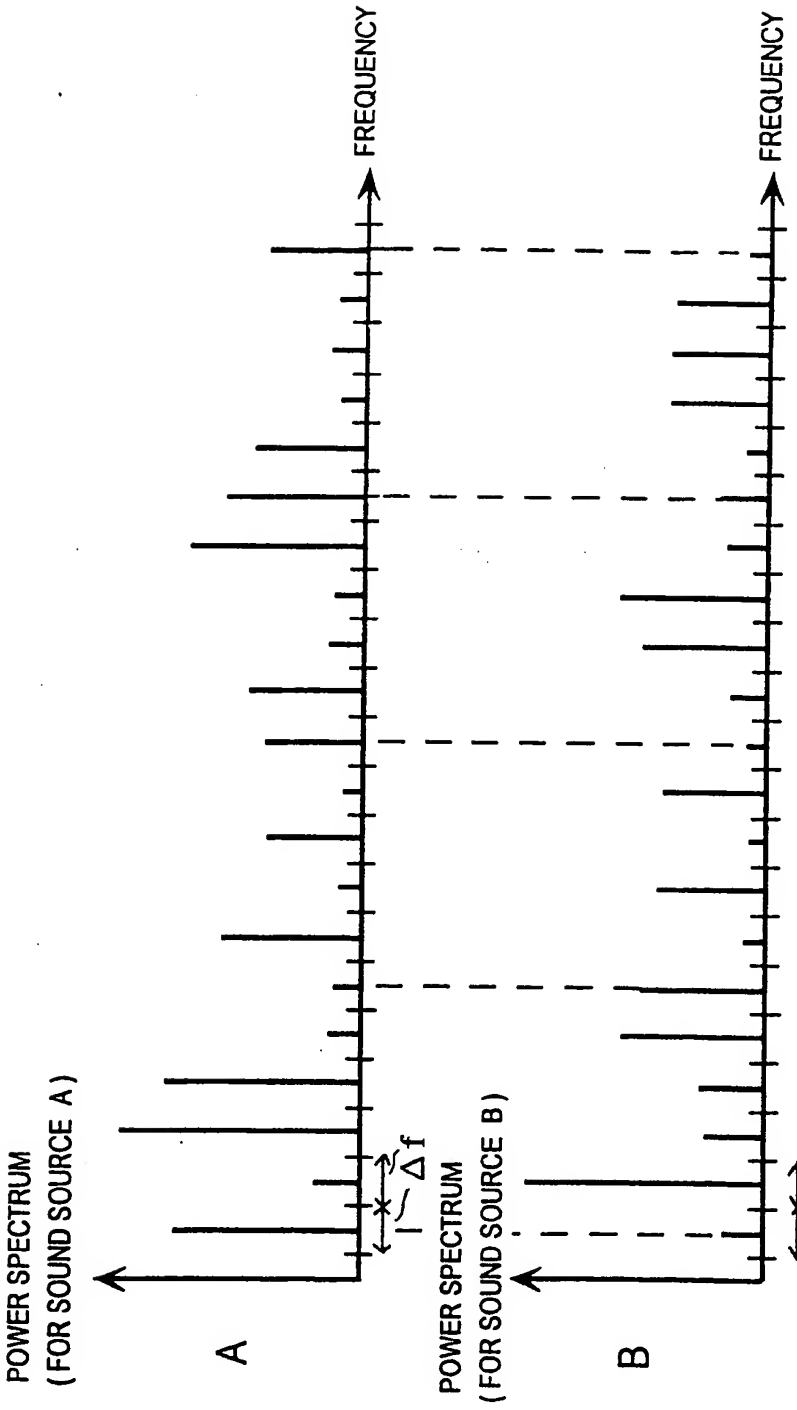



FIG. 4 EXAMPLES OF SPECTRUM STRUCTURES FOR SOUND SOURCES A, B



BNSDOCID: <EP_____0831458A2_I_>

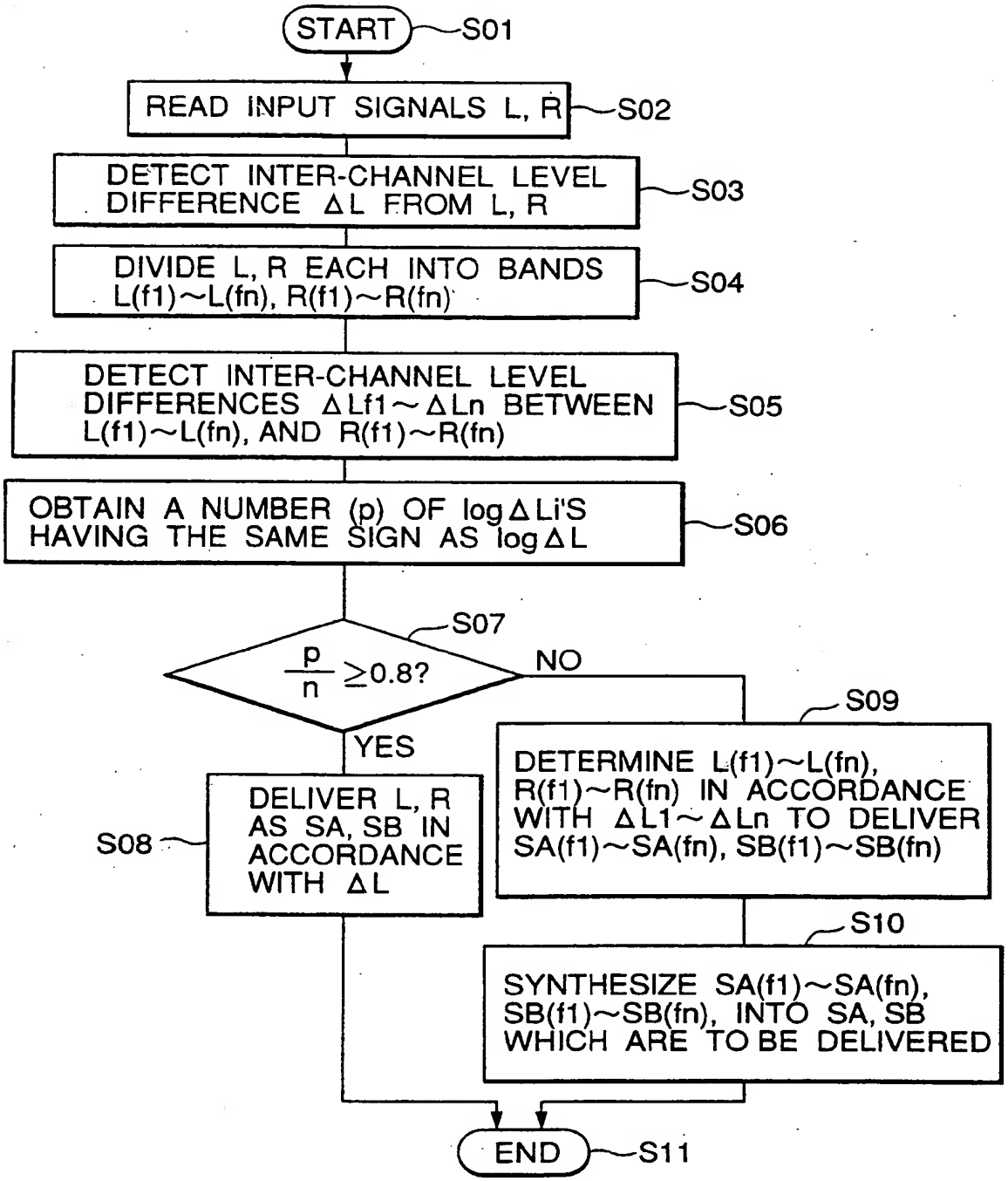



FIG.5



BNSDOCID: <EP_____0831458A2_I_>

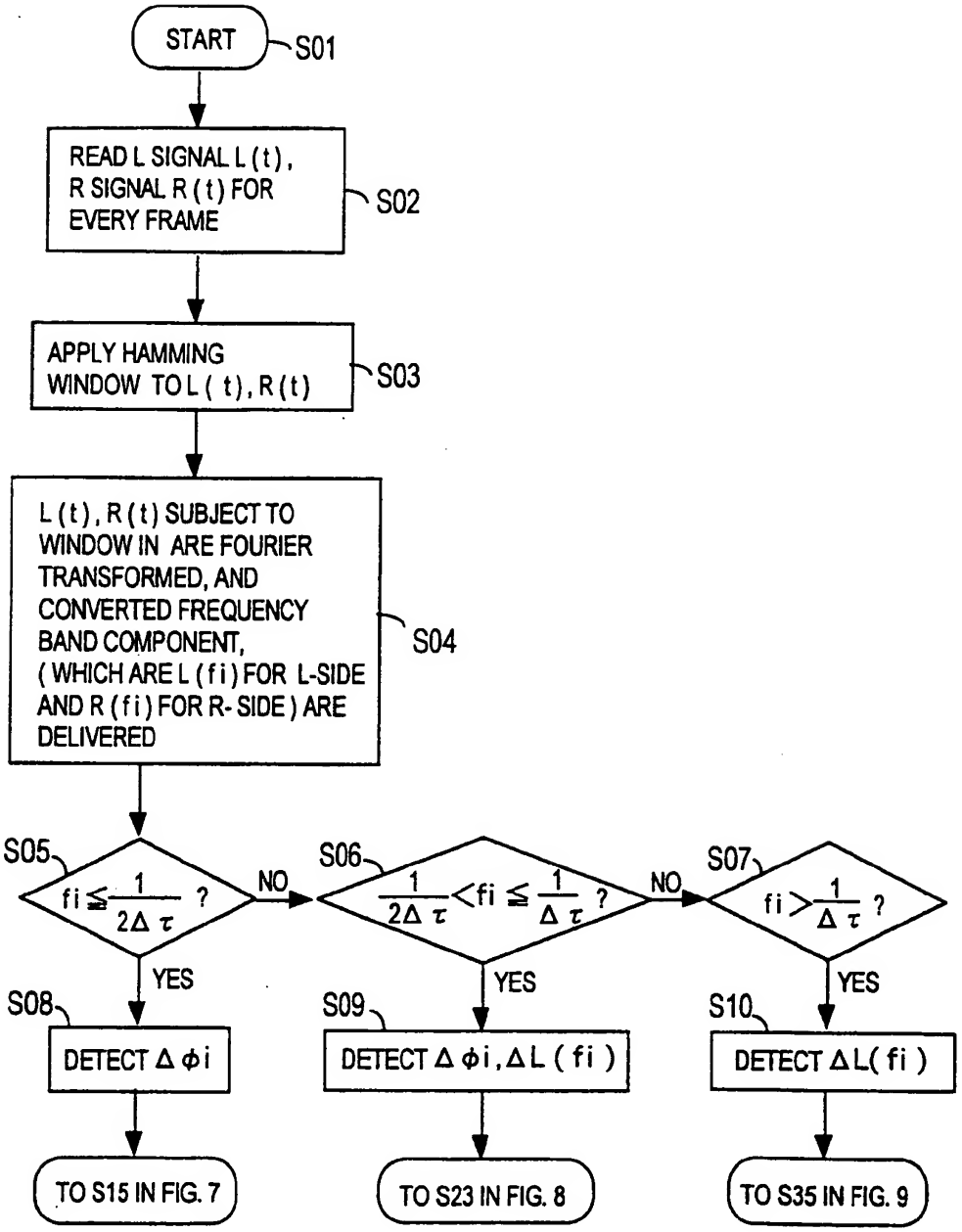



FIG. 6



BNSDOCID: <EP_____0831458A2_I_>

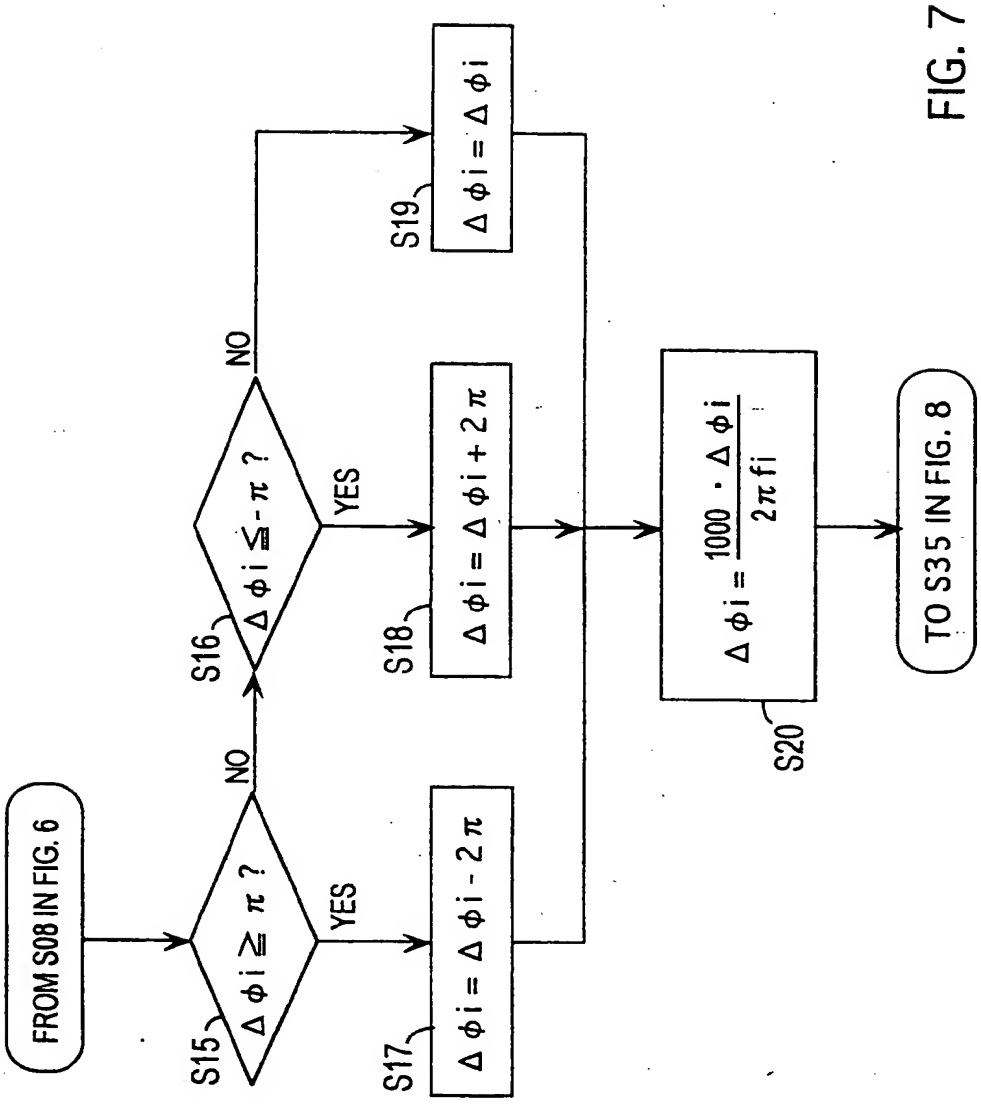



FIG. 7



BNSDOCID: <EP_____0831458A2_I_>

FOR $\frac{1}{2\Delta\tau} < fi \leqq \frac{1}{\Delta\tau}$

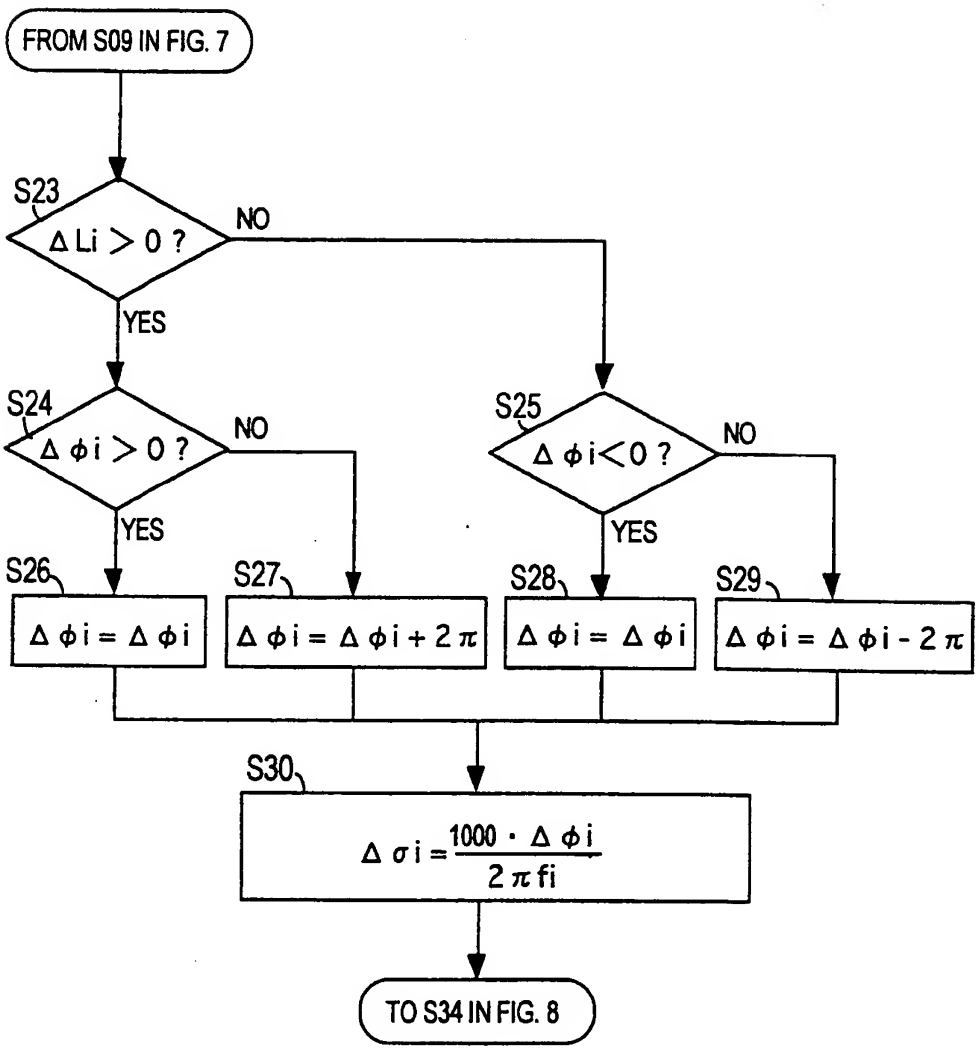



FIG. 8



BNSDOCID: <EP_____0831458A2_I_>

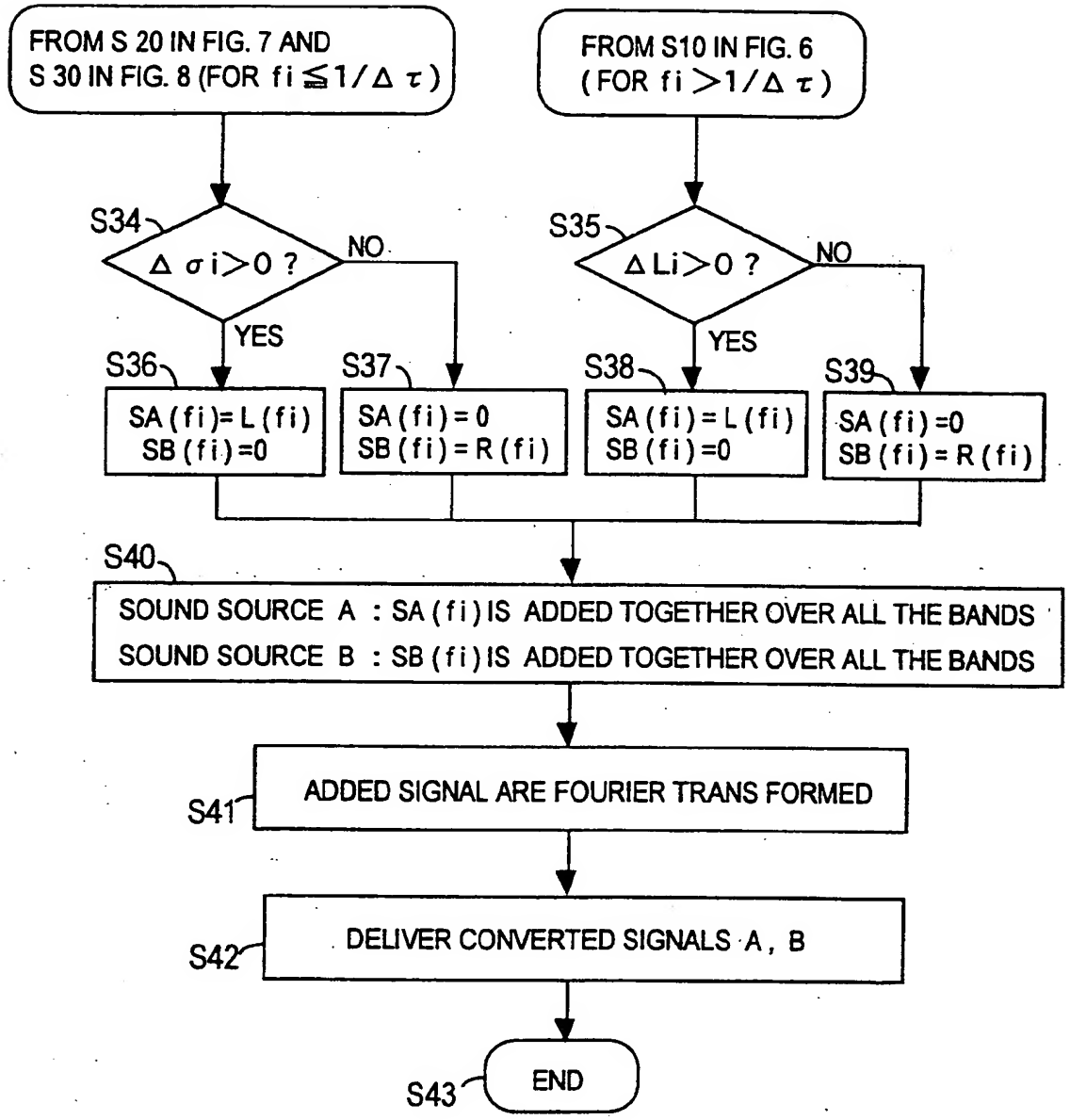



FIG. 9



BNSDOCID: <EP_____0831458A2_I_>

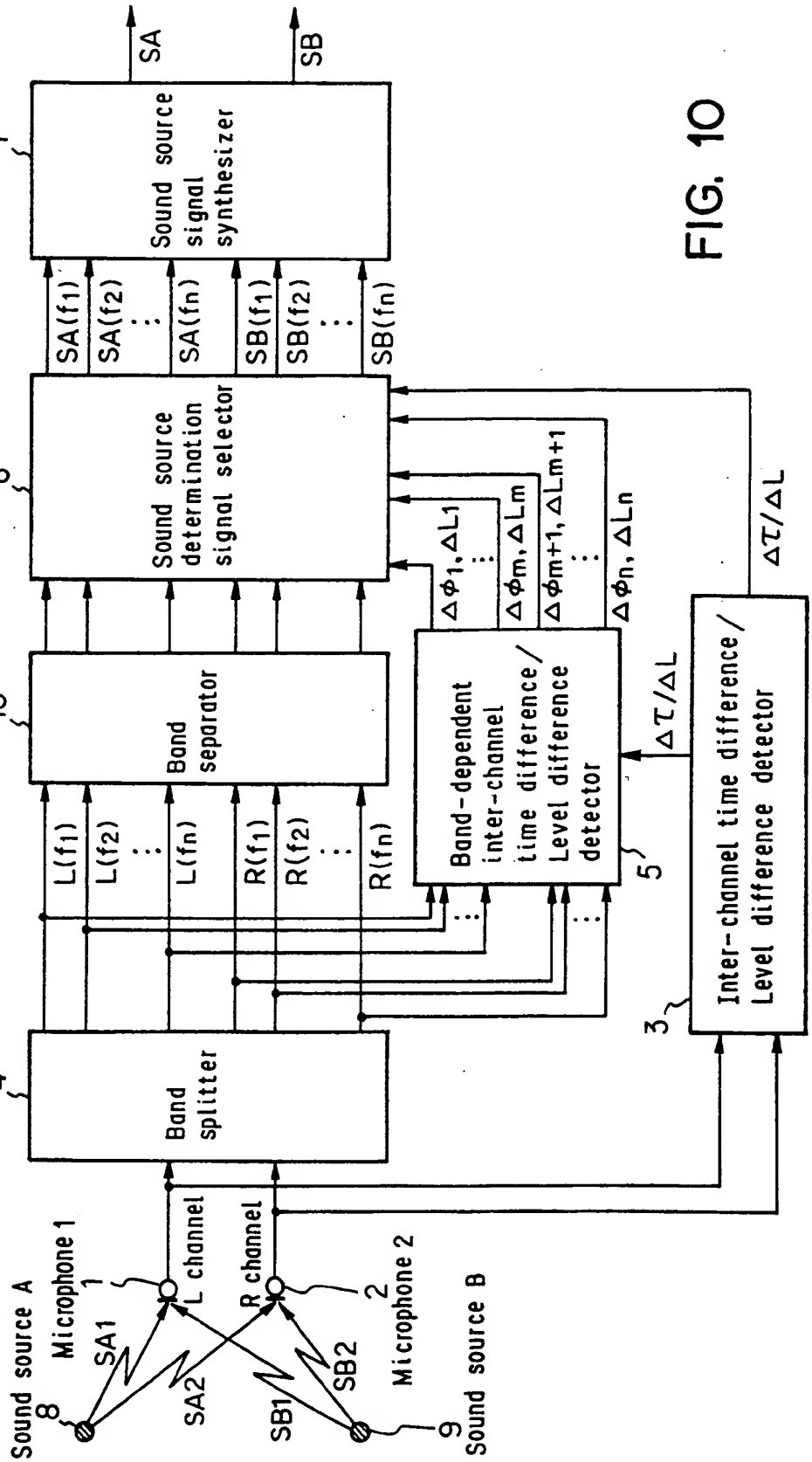



FIG. 10



BNSDOCID: <EP_____0831458A2_I_>

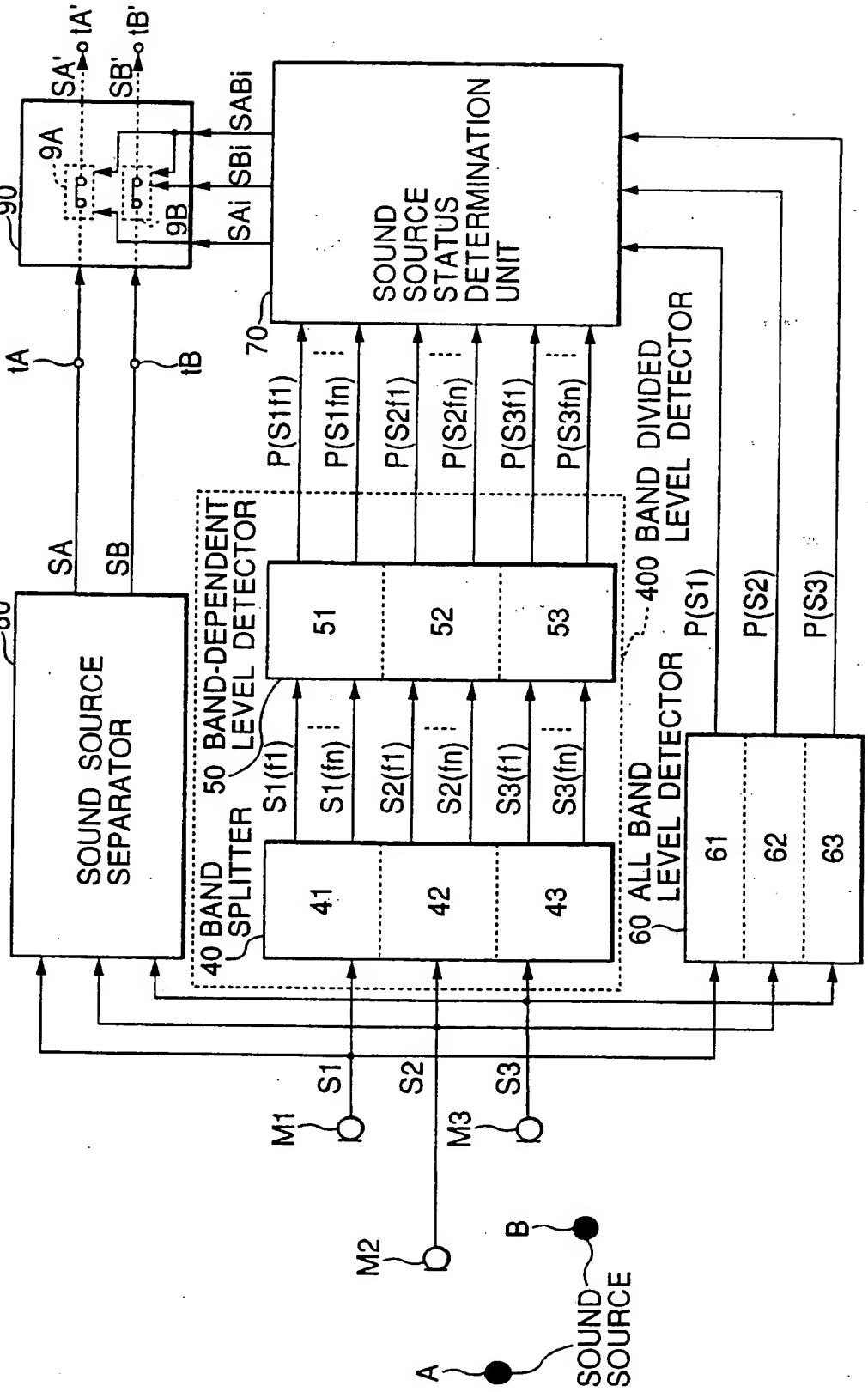



FIG.11



BNSDOCID: <EP_____0831458A2_I_>

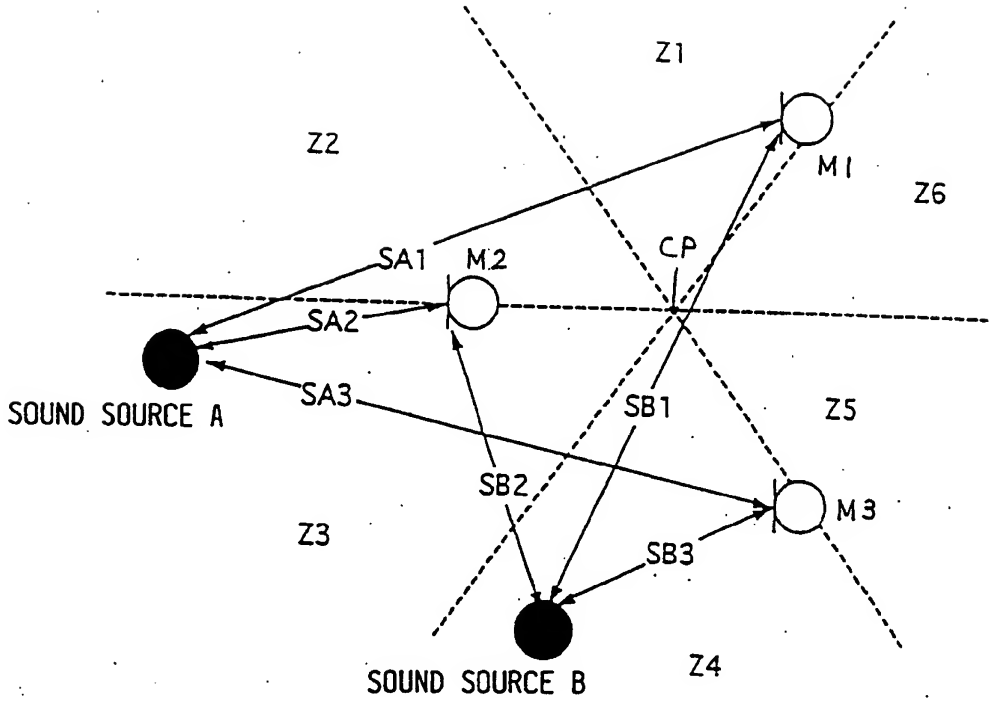



FIG.12



BNSDOCID: <EP_____0831458A2_I_>

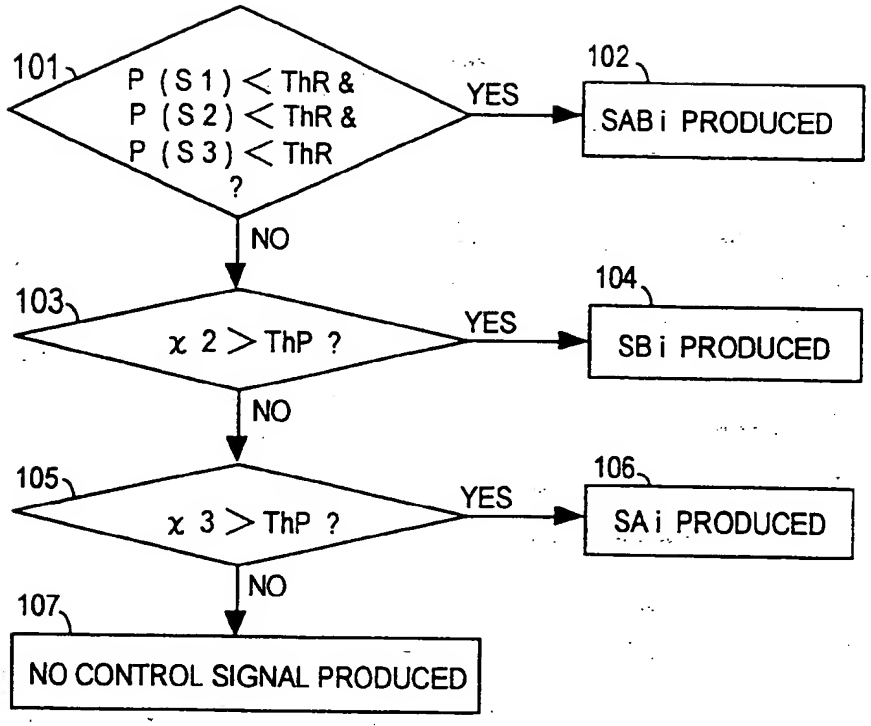



FIG. 13



BNSDOCID: <EP_____0831458A2_I_>

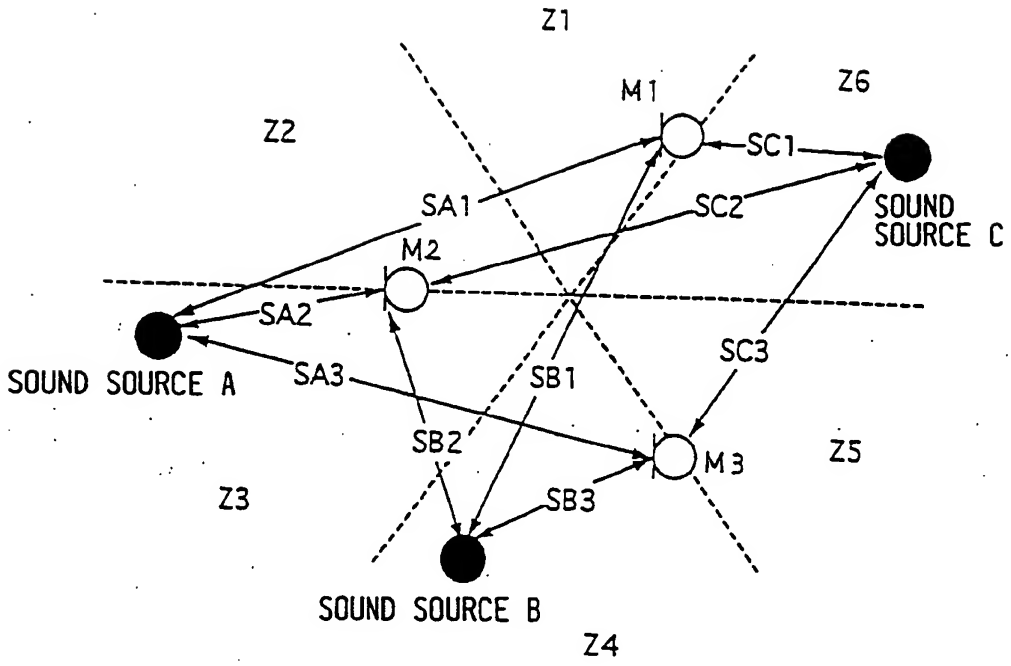



FIG.14



BNSDOCID: <EP_____0831458A2_I_>

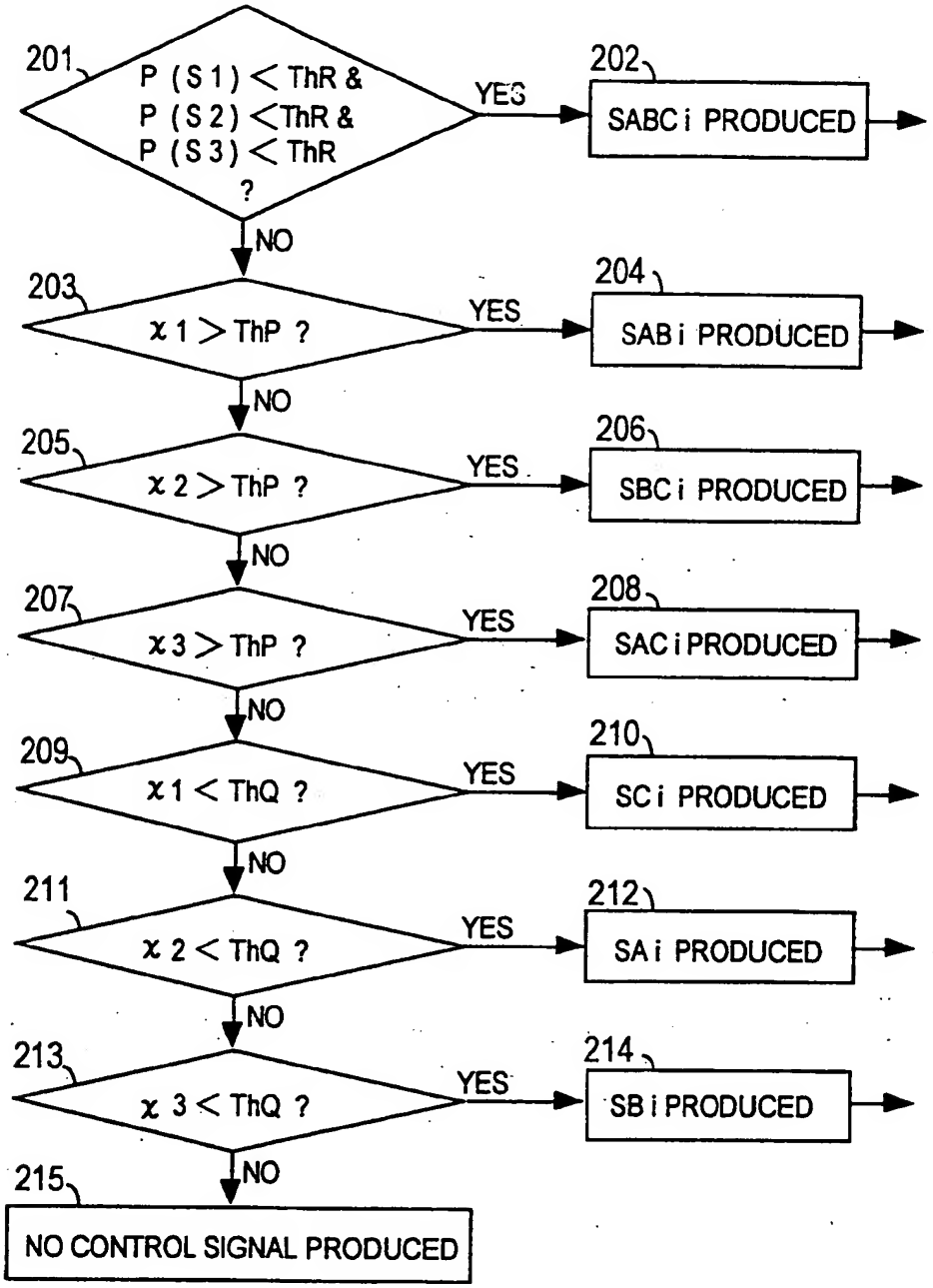



FIG. 15



BNSDOCID: <EP_____0831458A2_I_>

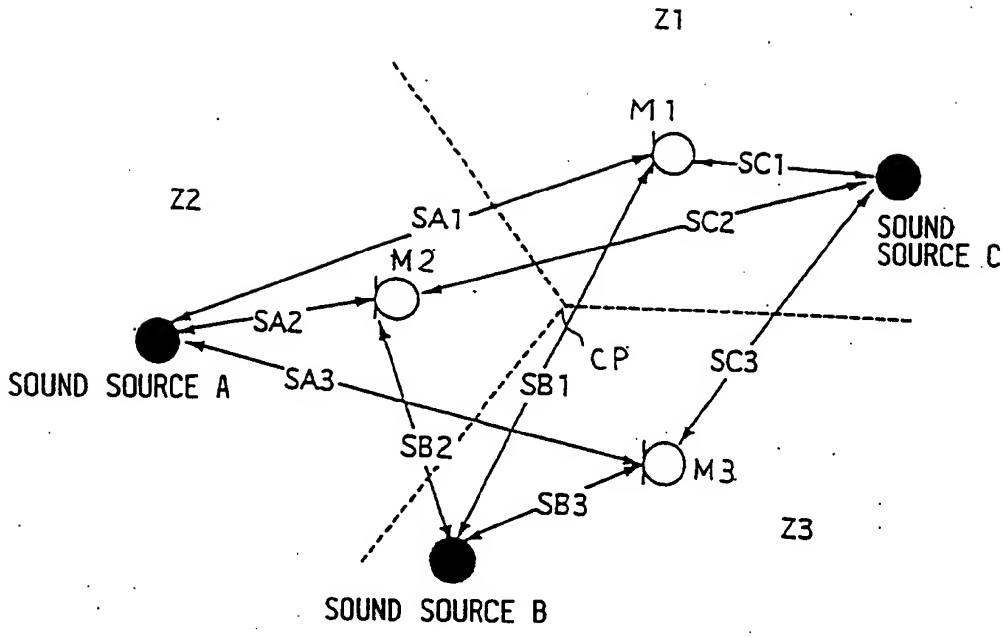



FIG.16



BNSDOCID: <EP_____0B31458A2_I_>

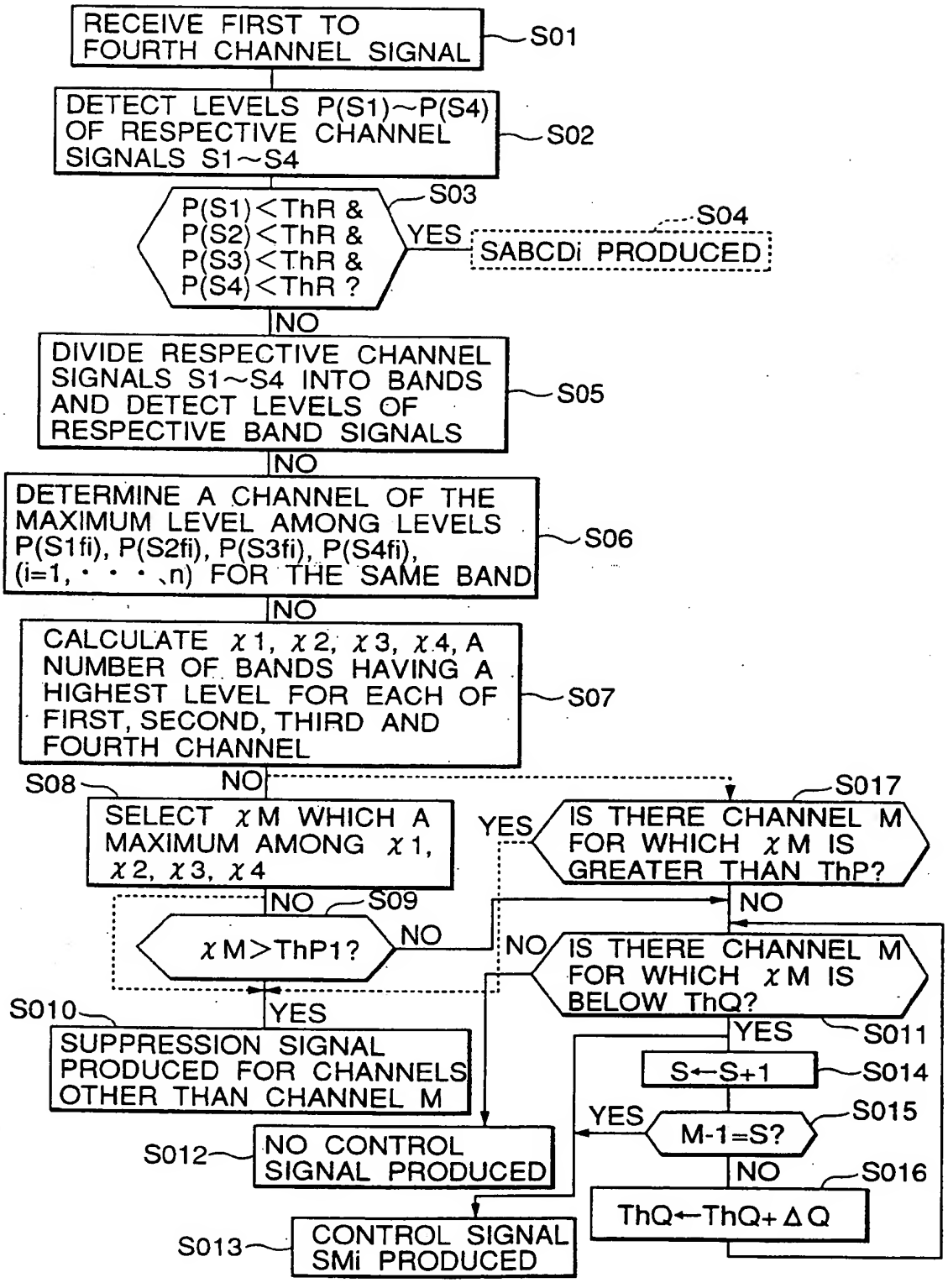



FIG.17



BNSDOCID: <EP_____0831458A2_I_>

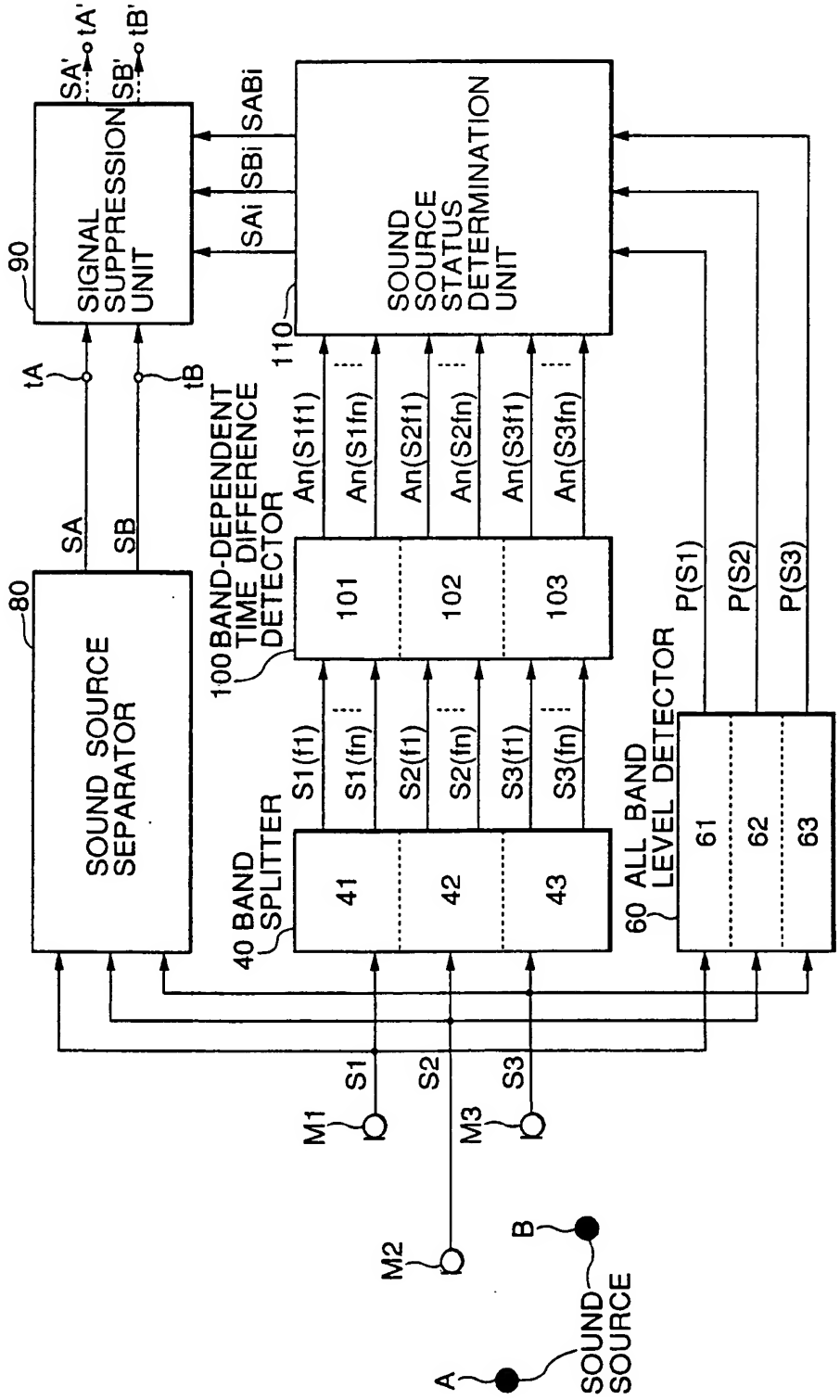



FIG.18



BNSDOCID: <EP_____0831458A2_I_>

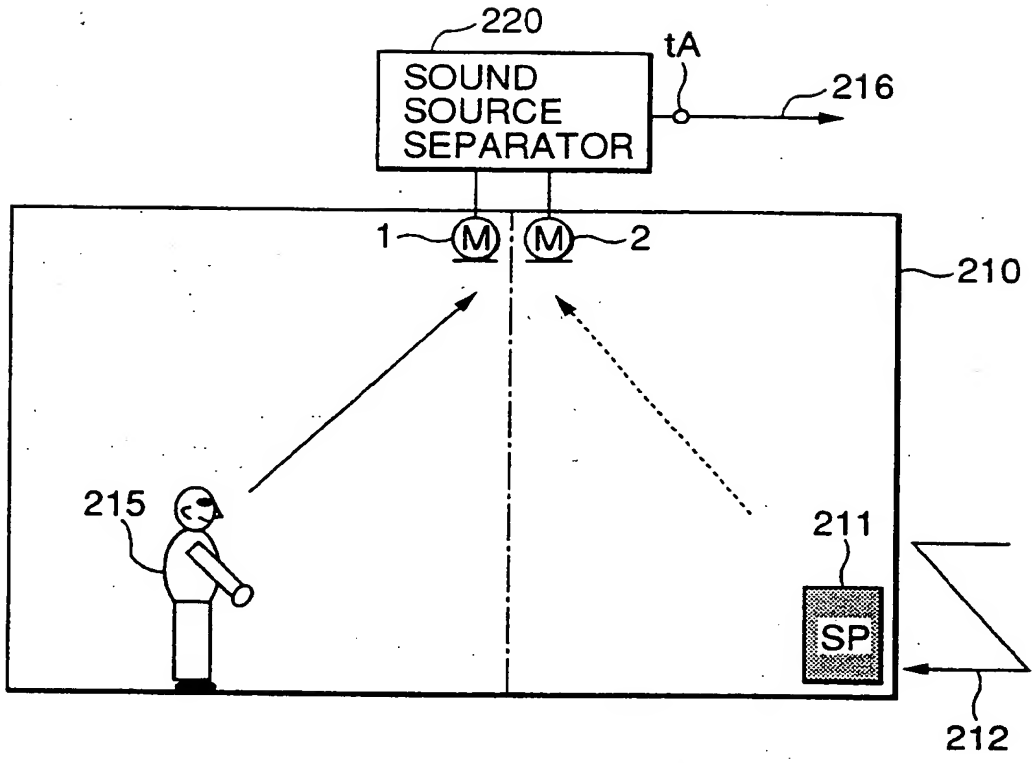



FIG.19



BNSDOCID: <EP_____0831458A2_I_>

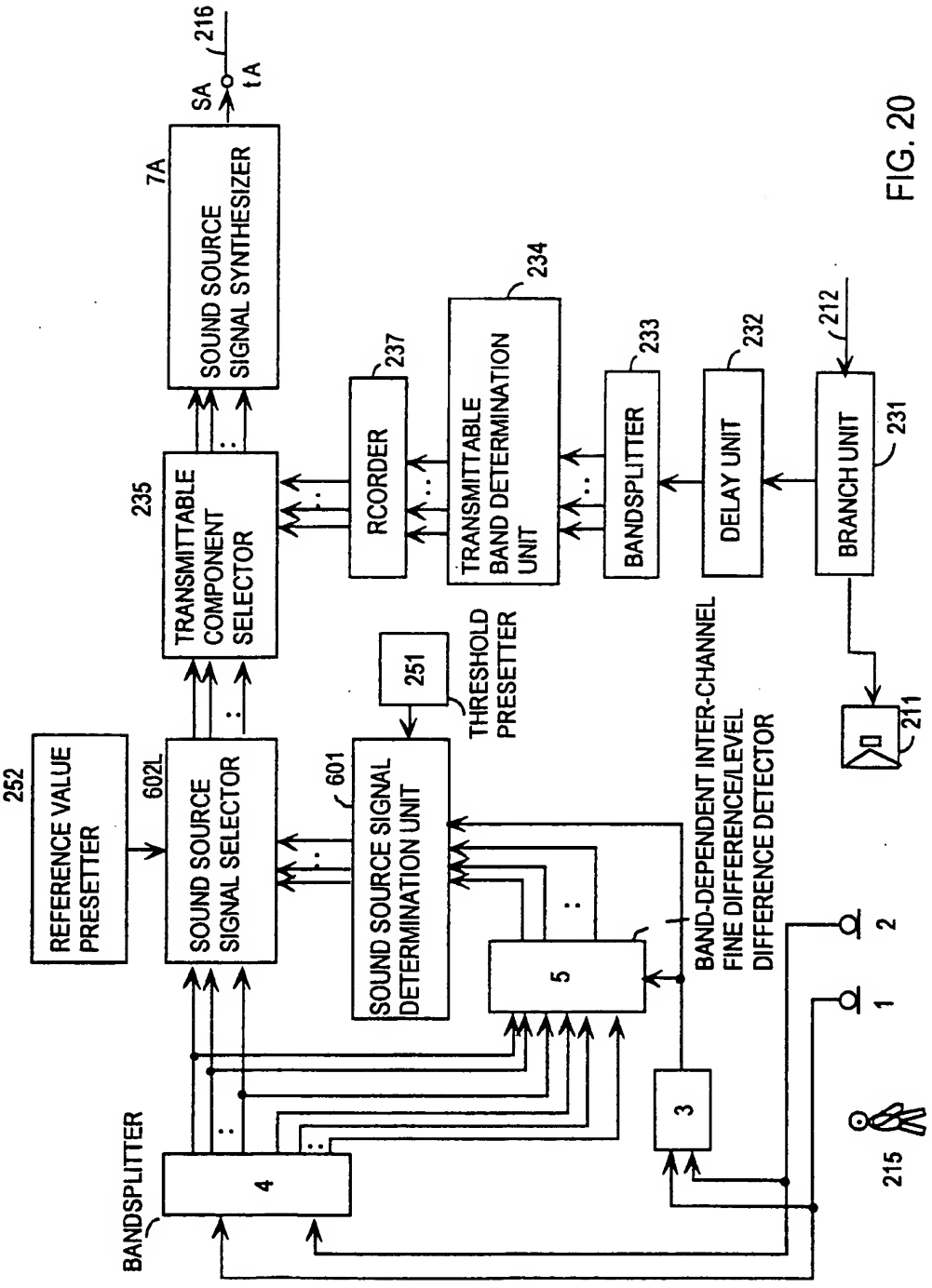



FIG. 20



BNSDOCID: <EP_____0831458A2_I_>

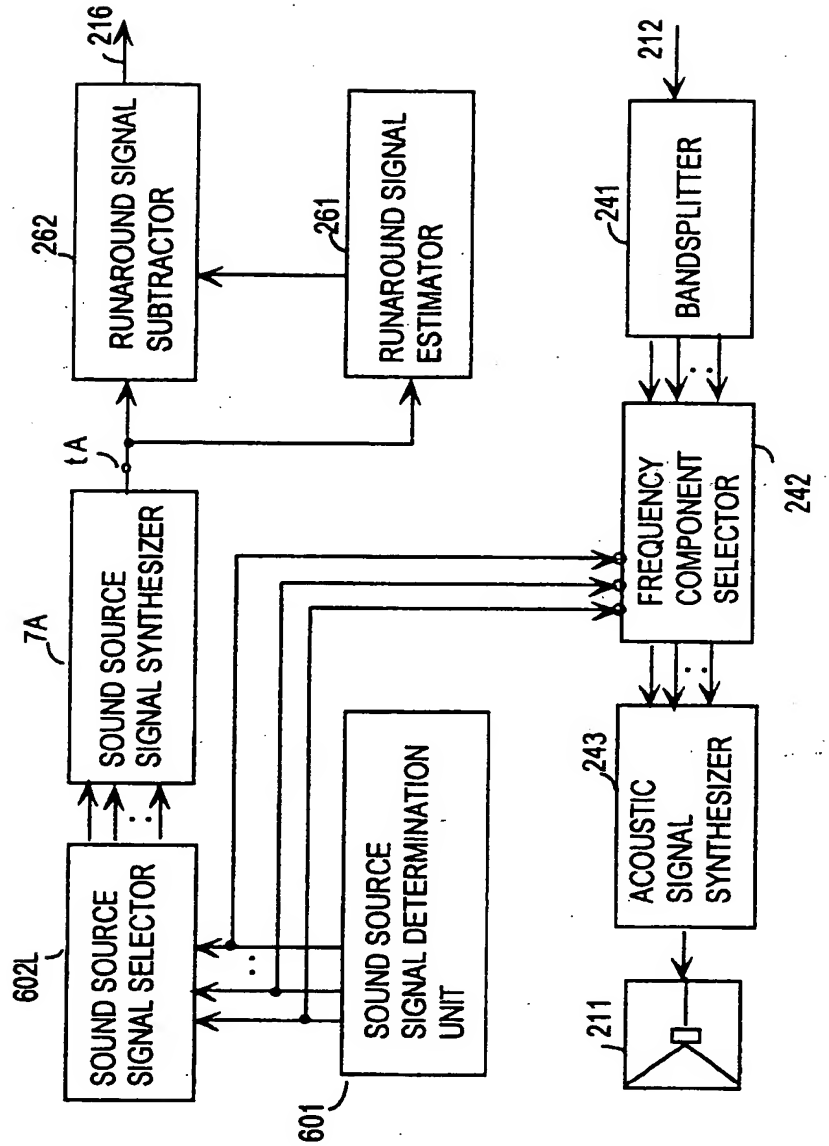



FIG. 21



BNSDOCID: <EP_____0831458A2_I_>

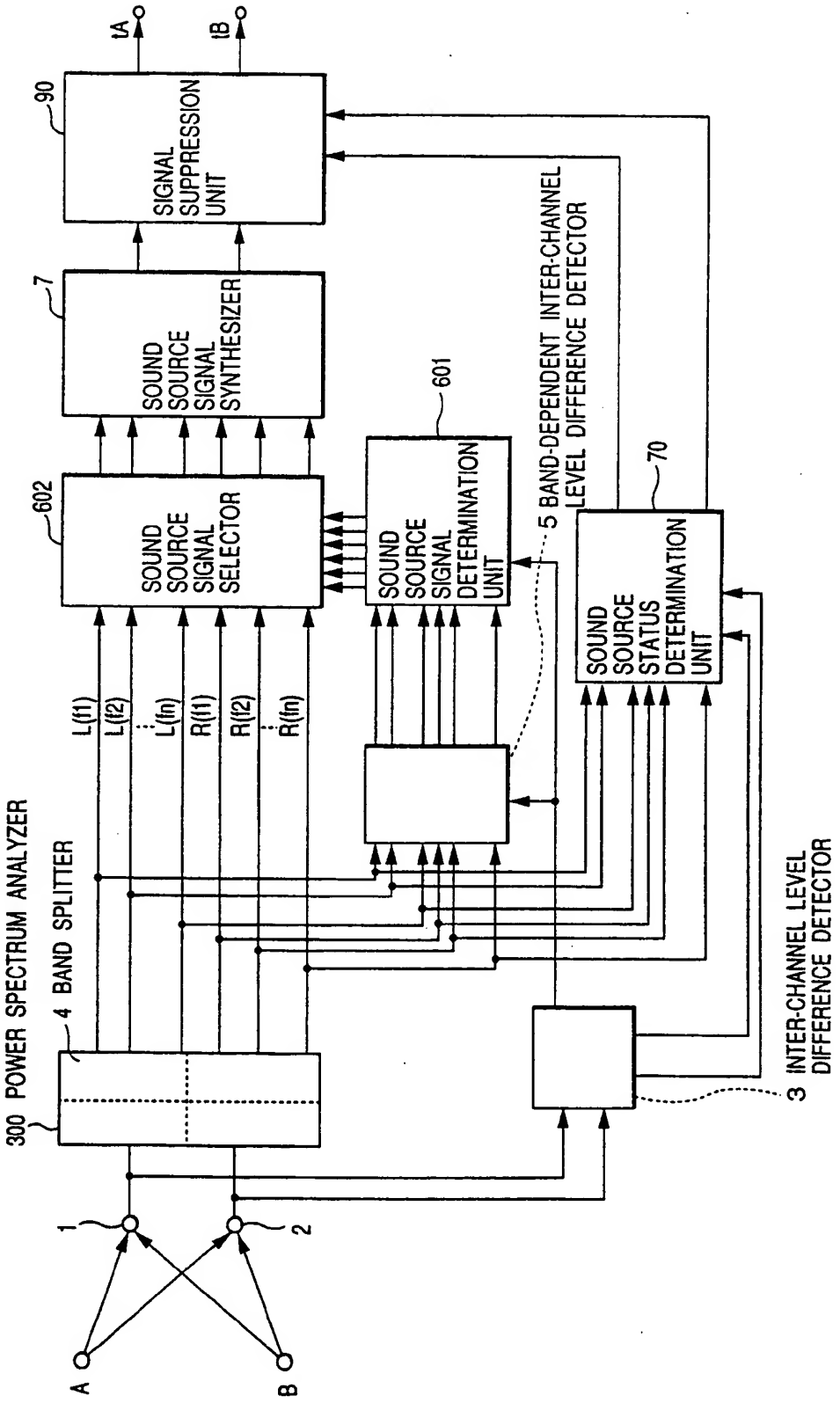



FIG.22



BNSDOCID: <EP_____0831458A2_I_>

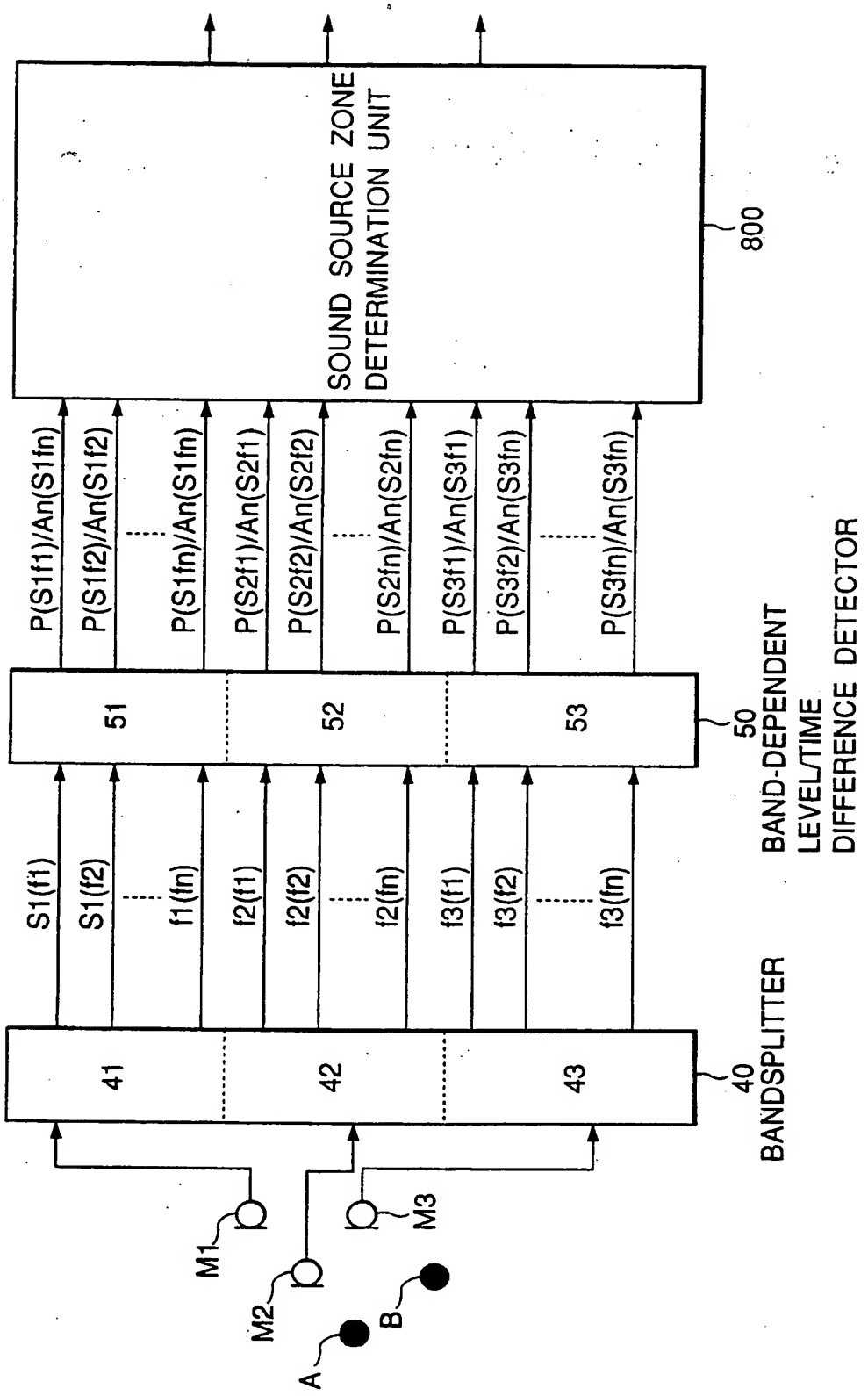



FIG.23



BNSDOCID: <EP_____0831458A2_I_>

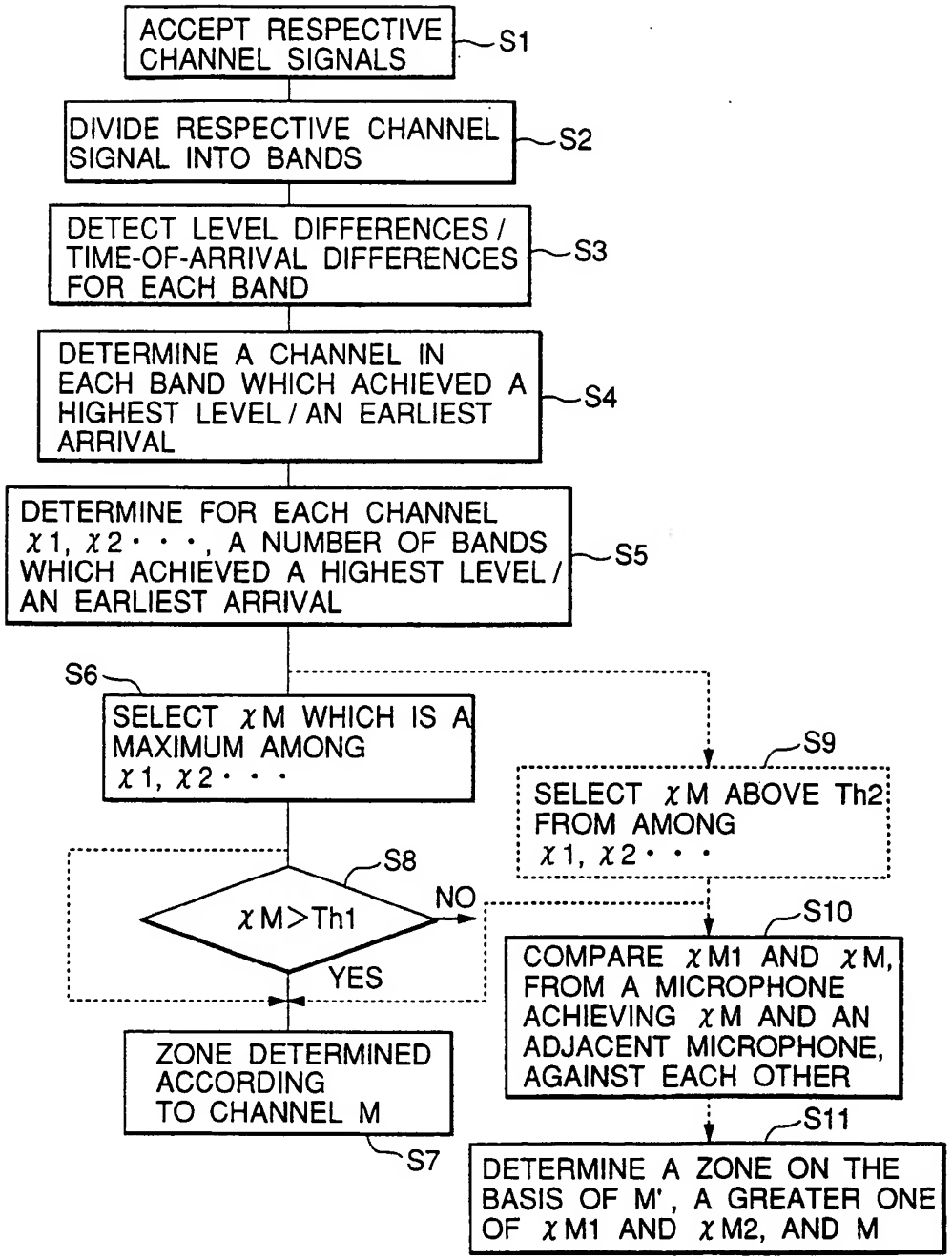



FIG.24



BNSDOCID: <EP_____0831458A2_I_>

| MIXED SIGNAL | SEPARATED SIGNAL |
|---|---|
| MALE VOICE+ FEMALE VOICE | MALE VOICE(S1) |
| | FEMALE VOICE(S2) |
| FEMALE VOICE1+ FEMALE VOICE2 (BY SAME SPEAKER, BUT SPEAKING DIFFERENT CONTENT) | FEMALE VOICE1(S3) |
| | FEMALE VOICE2(S4) |
| COCK CLUCKING+ FEMALE VOICE | COCK CLUCKING(S5) |
| | FEMALE VOICE(S6) |

FIG.25

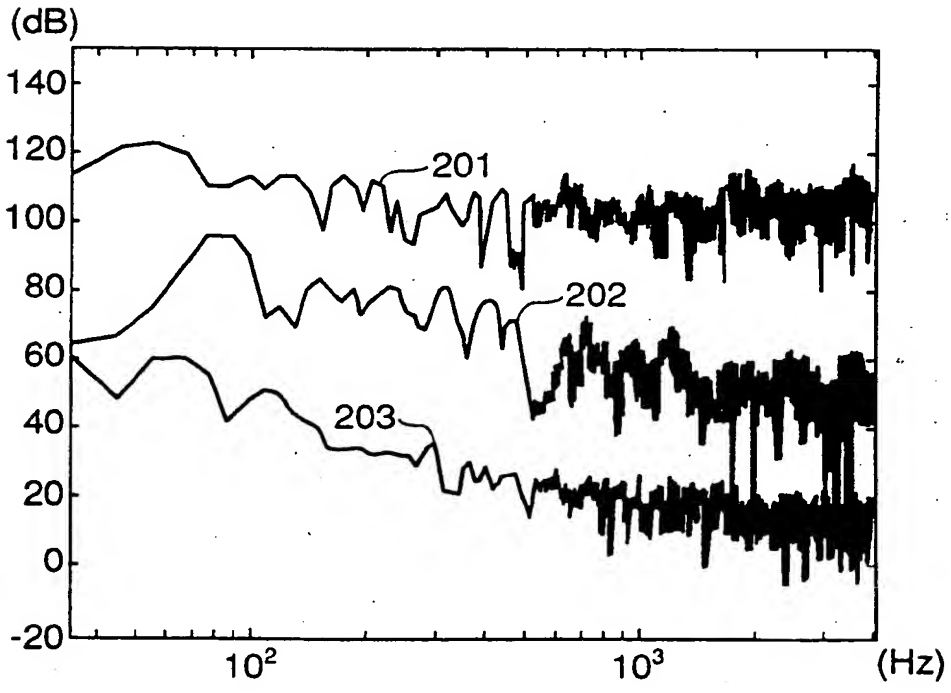



FIG.26



BNSDOCID: <EP_____0831458A2_I_>

FIG. 27A

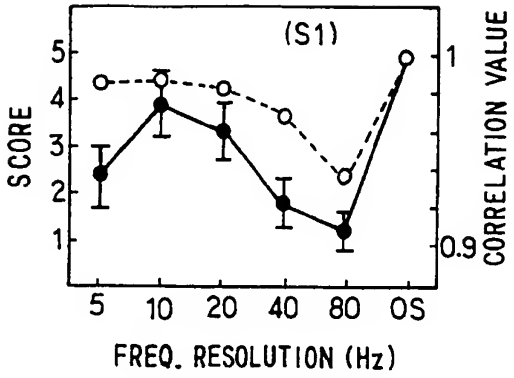



FIG. 27B

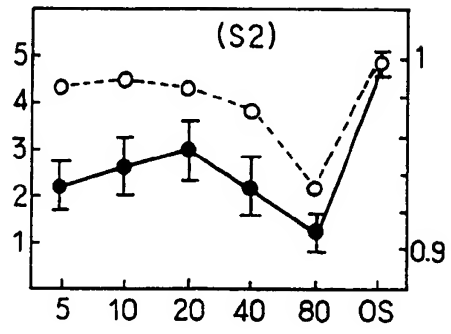



FIG. 27C

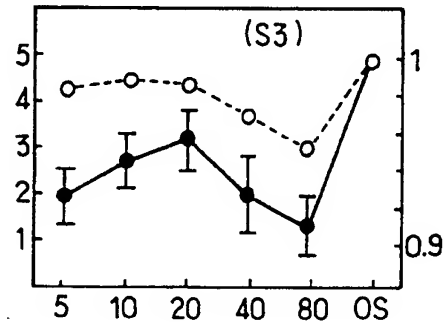



FIG. 27D

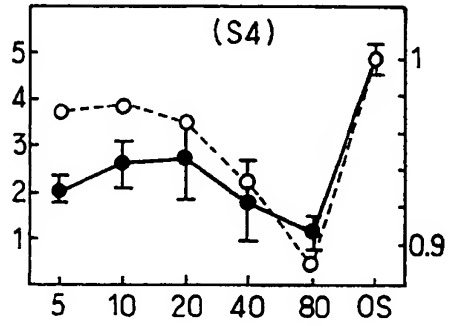



FIG. 27E

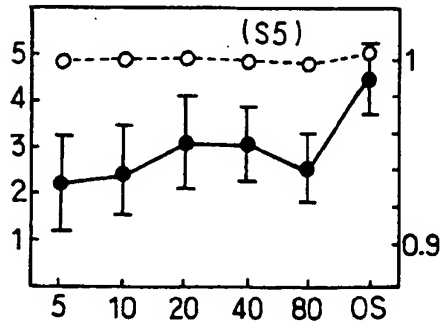
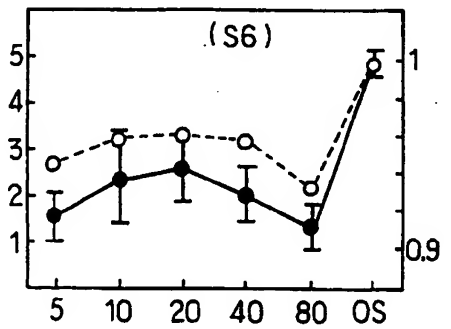



FIG. 27F



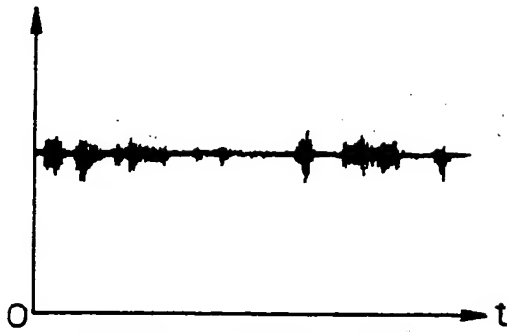



BNSDOCID: <EP_____0831458A2_I_>

FIG. 28A

ORIGINAL SIGNAL (S1)

FIG. 28B

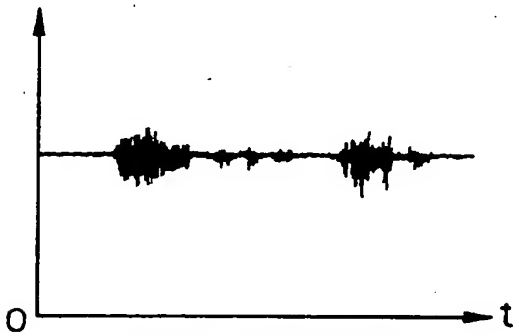



ORIGINAL SIGNAL (S2)

FIG. 28C

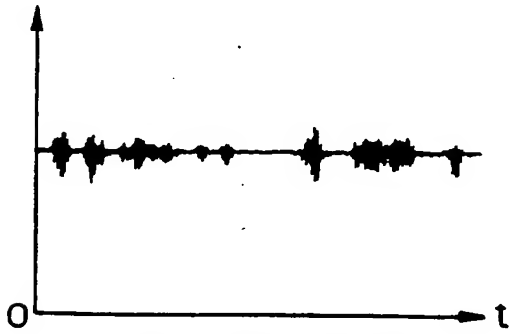
MIXED SIGNAL

FIG. 28D




POST-PROCESSING SIGNAL (S1)

FIG. 28E




POST-PROCESSING SIGNAL (S2)



BNSDOCID: <EP_____0831458A2_I_>

THIS PAGE BLANK (USPTO)

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 831 458 A3

# (12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
**11.11.1998 Bulletin 1998/46**

(43) Date of publication A2:
**25.03.1998 Bulletin 1998/13**

(21) Application number: **97116245.8**

(22) Date of filing: **18.09.1997**

(51) Int. Cl.[6]: **G10L 5/02**, G10L 7/02, G10L 9/06, G10L 9/18, G10L 3/00, G10K 11/16

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE**
Designated Extension States:
**AL LT LV RO SI**

(30) Priority: **18.09.1996 JP 246726/96**
**13.03.1997 JP 76668/97**
**13.03.1997 JP 76672/97**
**13.03.1997 JP 76682/97**
**13.03.1997 JP 76693/97**
**13.03.1997 JP 76695/97**

(71) Applicant:
**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**
**Shinjuku-ku, Tokyo 163-19 (JP)**

(72) Inventors:
- **Aoki, Mariko**
  **Yokohama-shi, Kanagawa 227 (JP)**
- **Aoki, Shigeaki**
  **Yokosuka-shi, Kanagawa 238 (JP)**
- **Matsui, Hiroyuki**
  **Yokohama-shi, Kanagawa 235 (JP)**
- **Nishino, Yutaka**
  **Miura-shi, Kanagawa 238-01 (JP)**
- **Okamoto, Manabu**
  **Yokohama-shi, Kanagawa 236 (JP)**

(74) Representative:
**Hoffmann, Eckart, Dipl.-Ing.**
**Patentanwalt,**
**Bahnhofstrasse 103**
**82166 Gräfelfing (DE)**

## (54) Method and apparatus for separation of sound source, program recorded medium therefor, method and apparatus for detection of sound source zone; and program recorded medium therefor

(57) A time difference $\Delta\tau$ between the arrival of acoustic signals from sound sources to microphones 1, 2 is detected from output channel signals L, R from microphones 1, 2. By Fourier transform, the signals L, R are divided into respective frequency bands L(f1) - L(fn), R(f1) - R(fn). Differences $\Delta\tau_i$ ( i = 1, 2, • • • n ) in the time-of-arrival of L(f1) - L(fn) and R(f1) - R(fn) to the microphones 1, 2 as well as a signal level difference $\Delta$Li are detected. L(f1) - L(fn), R(f1) - R(fn) are divided into a low range of $fi < 1/(2\Delta\tau)$, a middle range of $1/(2\Delta\tau) < fi < 1/\Delta\tau$, and a high range of $fi > 1/\Delta\tau$. Utilizing $\Delta\tau_i$ for the low range, $\Delta$Li and $\Delta\tau_i$ for the middle range and $\Delta$Li for the high range, a determination is made from which sound source L(fi), R(fi) are oncoming to deliver outputs separately for each sound source. The outputs are subject to an inverse Fourier transform for synthesis separately for each sound source.

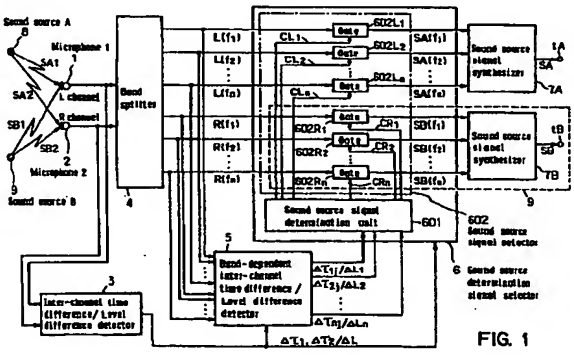


FIG. 1

EP 0 831 458 A3



BNSDOCID: <EP_____0831458A3_I_>

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number

## DOCUMENTS CONSIDERED TO BE RELEVANT

EP 97116245.8

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int. Cl. 6) |
|---|---|---|---|
| A,P | EP 0795851 A2 (TOSHIBA K.K.) 17 September 1997 (17.09.97), abstract, fig. 4, claim 1, page 3, line 17 - page 4, line 29. -- | 1 | G 10 L 5/02<br>G 10 L 7/02<br>G 10 L 9/06<br>G 10 L 9/18<br>G 10 L 3/00<br>G 10 K 11/16 |
| A | GB 2276298 A (CENTRAL RESEARCH LABORATORIES LTD.) 21 September 1994 (21.09.94), abstract, fig. 1,2,4,5, claims 1-11. -- | 1 | |
| A | EP 0509654 A2 (HEWLETT-PACKARD COMP.) 21 October 1992 (21.10.92), abstract, fig. 1,5, claims 1-10. -- | 1 | |
| A | WO 94/22278 A1 (CENTRAL RESEARCH LABORATORIES LTD.) 29 September 1994 (29.09.94), abstract, fig. 1,2,4,5, claims 1-14. -- | 1 | TECHNICAL FIELDS SEARCHED (Int. Cl. 6)<br>G 10 L 3/00<br>G 10 L 5/00<br>G 10 L 7/00<br>G 10 L 9/00 |
| A | GB 2275388 A (FUJI JUKOGYO K.K.) 24 August 1994 (24.08.94), fig. 1, abstract, claim 1. ---- | 1 | H 04 S 1/00<br>H 04 S 5/00<br>H 04 R 5/00<br>H 04 R 29/00<br>G 10 K 11/00 |

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| VIENNA | 27-07-1998 | BERGER |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)



BNSDOCID: <EP_____0831458A3_I_>